DEVELOPMENT AND EVALUATION OF AN AUTOMATED PATH PLANNING AID

A Thesis
Presented to
The Academic Faculty

by

Robert M. Watts

In Partial Fulfillment
of the Requirements for the Degree
Master of Science in the
School of Aerospace Engineering

Georgia Institute of Technology
May 2010

DEVELOPMENT AND EVALUATION OF AN AUTOMATED PATH PLANNING AID

Approved by:

Lockheed Martin Associate Professor of
Avionics Integration Eric Johnson, Advisor
School of Aerospace Engineering
Georgia Institute of Technology

Professor Panagiotis Tsiotras
School of Aerospace Engineering
Georgia Institute of Technology

Assistant Professor Karen Feigh
School of Aerospace Engineering
Georgia Institute of Technology

David S. Lewis Associate Professor of
Cognitive Engineering Amy Pritchett
School of Aerospace Engineering
Georgia Institute of Technology

Date Approved: 5 April 2010

ACKNOWLEDGEMENTS

I would first like to thank each of the committee members. Dr. Panagiotis Tsiotras provided excellent feedback on every aspect of the design of this project. The experiment itself could not have been done without the patient advice of Dr. Karen Feigh. And special thanks go to my advisor, Dr. Eric Johnson, for his knowledge and guidance in helping to complete the project.

I would also like to thank my labmates who provided their input and encouragement along the way. I also owe a special thanks to Zarrin Chua and So Young Kim for their help, advice and encouragement in accomplishing a human subjects experiment. I need to also thank Corey Ippolito and Ted Chen for their development of the Reconfigurable Flight Simulator, on which this prototype was built. Though I have never met them, without their work, the prototype development would have taken years instead of months.

Thank you to Mr. David Pietruszewski who provided knowledge and expertise from an airline pilot's perspective. Many aspects of this project could not have happened without his advice. The pilots who participated in this experiment also deserve much of my thanks. Each of them took the time to come in and test the prototype and provided so much more insight than simple results could have given. I thank each of them for the interest they took in the success of this project.

Finally, and most importantly I would like to thank my family, who have always supported me in everything I have done. They provided me support when I needed it and also distractions when that was what I needed most. I love you and thank you.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

SUMMARY

In the event of an onboard emergency, air transport pilots are remarkably adept at safely landing their aircraft. However, the tasks of selecting an alternate landing site and developing a safe path to land are very difficult in the high workload, high stress environment of a cockpit during an emergency. The purpose of this research was to develop an automated path planning aid which would assist the pilot in the completion of these tasks. A prototype was developed to test this concept experimentally. The experiment was also intended to gather further information about how pilots think about and accomplish this task as well as the best ways to assist them.

In order to better understand the priorities and processes pilots use when dealing with emergency planning, a survey of airline pilots was conducted. The results of this survey highlighted the fact that each emergency is unique and has its own set of factors which are critically important. One factor which is important in many emergencies is the need to land quickly. The survey responses indicated that one of the most important characteristics of a useful tool is that it should provide pertinent information in an easy to use manner, and should not divert too much attention from their other tasks.

A number of design goals drove the development of the prototype aid. First, the aid was to work within current aircraft, without requiring substantial redesign on the cockpit. Second, the aid was to help improve pilots' performance without increasing their workload. Finally, the aid was designed to assist pilots in obtaining and processing critical information which influences the site selection and path development tasks. One variation of the aid included a filter dial which allowed pilots to quickly

reduce the number of options considered, another variation of the aid did not include such a dial. These two variations of the aid were tested in order to assess the impact of the addition of the filter dial to the system.

Though many of the results did not prove to be statistically significant, they suggest that the addition of a filter dial improved the quality of the selected landing site; however, it also increased the time required for the selection. The results were obtained in both familiar and unfamiliar emergencies. The dial was shown to improve the time to complete the task in the case of unfamiliar emergencies. The experiment also compared an optimal ranking system to a non-optimal system, for which results showed no significant difference between the two. This may imply that while pilots did not tend to over rely on the ranking system, under-reliance may need to be addressed by training and a better understanding of the factors which impact the rankings.

The participants found that the aid facilitates quick and easy access to critical information. The aid was also useful for processing this information by filtering out options which were inappropriate for a given scenario through the use of the filter dial. The participants also made recommendations about possible improvements which could be made to the system such as better filter settings which are more similar to the way that pilots think about their options.

CHAPTER I

INTRODUCTION

Modern air transportation has a very good record of flight safety. When failures do occur in flight, the training and experience of the pilots almost always provide for a safe landing. This is evidenced by a rate of only 1.35 accidents per one million hours flown in 2007 by US air carriers [26]. Despite this excellent record, the pilots' responsibility to land safely in case of an emergency can be very demanding. When an emergency situation occurs during a flight, the pilots' workload is very high and a number of tasks demand the pilots' attention. Among these are the planning and execution of a trajectory that will result in a safe landing. However, this task is complicated by multiple, often conflicting goals, including reducing time in the air, staying within flight envelope limits as well as meeting regulatory requirements. Moreover, all these tasks must be accomplished in a stressful environment under time pressure [6].

In order to plan a safe landing, the pilots must first determine the best landing site. Then they must formulate an efficient and safe trajectory to the ground. The purpose of this research was to develop an Automated Planning Aid (APA) which assists pilots in the process of generating possible alternative landing sites, and evaluating them to choose the best option. Commercial airline pilots were recruited to participate in a number of simulated flights using the APA. The results were evaluated to determine the utility of such an aid.

1.1 Motivation

Between 1999 and 2008 there were a total of 91 fatal accidents on board commercial jet aircraft worldwide [2]. Figure 1 shows the number of fatal accidents by category, with many of the accidents categorized as either “Loss of Control - In flight” or “Controlled Flight into or Toward Terrain.” Some of these incidents may not have become fatal accidents if the pilots had been able to quickly plan and execute a satisfactory trajectory in order to complete a safe landing.

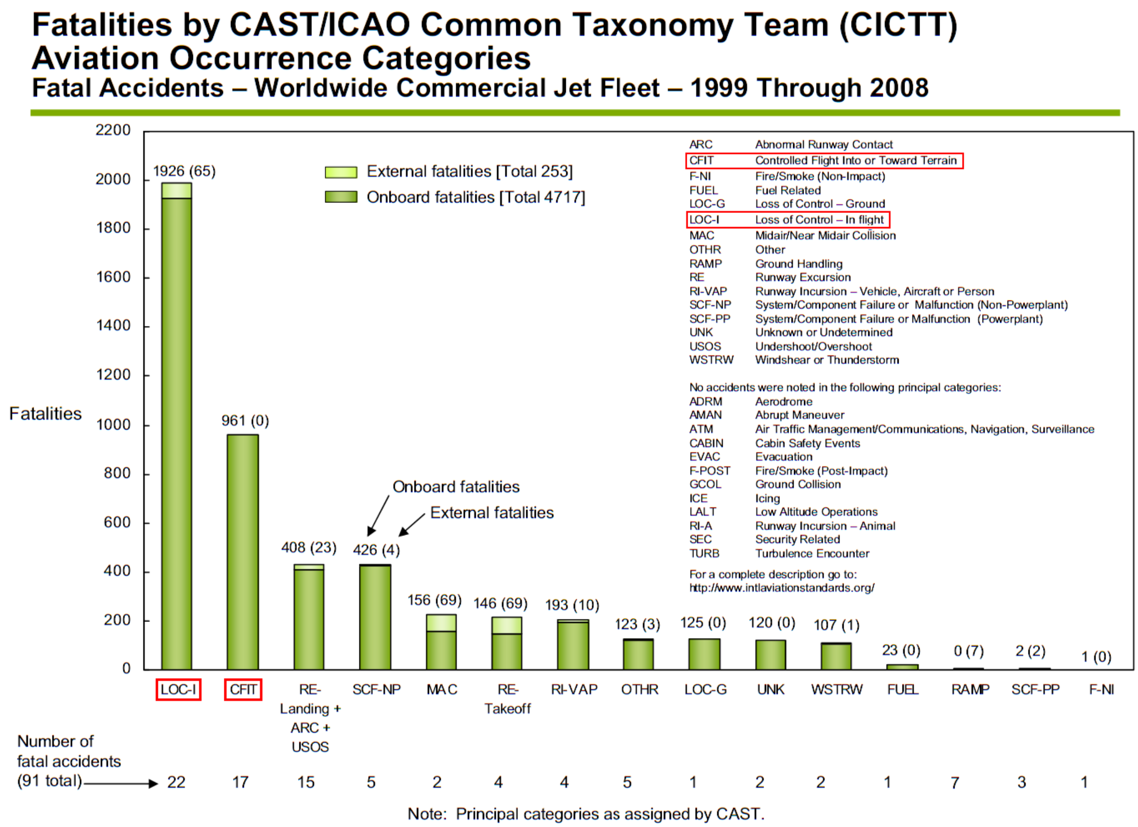


Figure 1: Fatal Accidents by Category, Commercial Jets 1999-2008 [2].

One example in which the crew could have benefited from the an automated planning aid is the crash of Swissair flight 111, which encountered smoke in the cockpit during its flight from New York to Geneva. When the pilots noticed the smoke, they declared an emergency (see Figure 2). After making an initial turn toward Boston, the controller recommended that they divert to Halifax instead. It was four minutes later when the pilots received the Halifax approach plates and realized that they were too high and needed to lose altitude. So they decided to circle around and dump fuel near Peggy’s Cove, Nova Scotia. After the aircraft had turned away from the airport

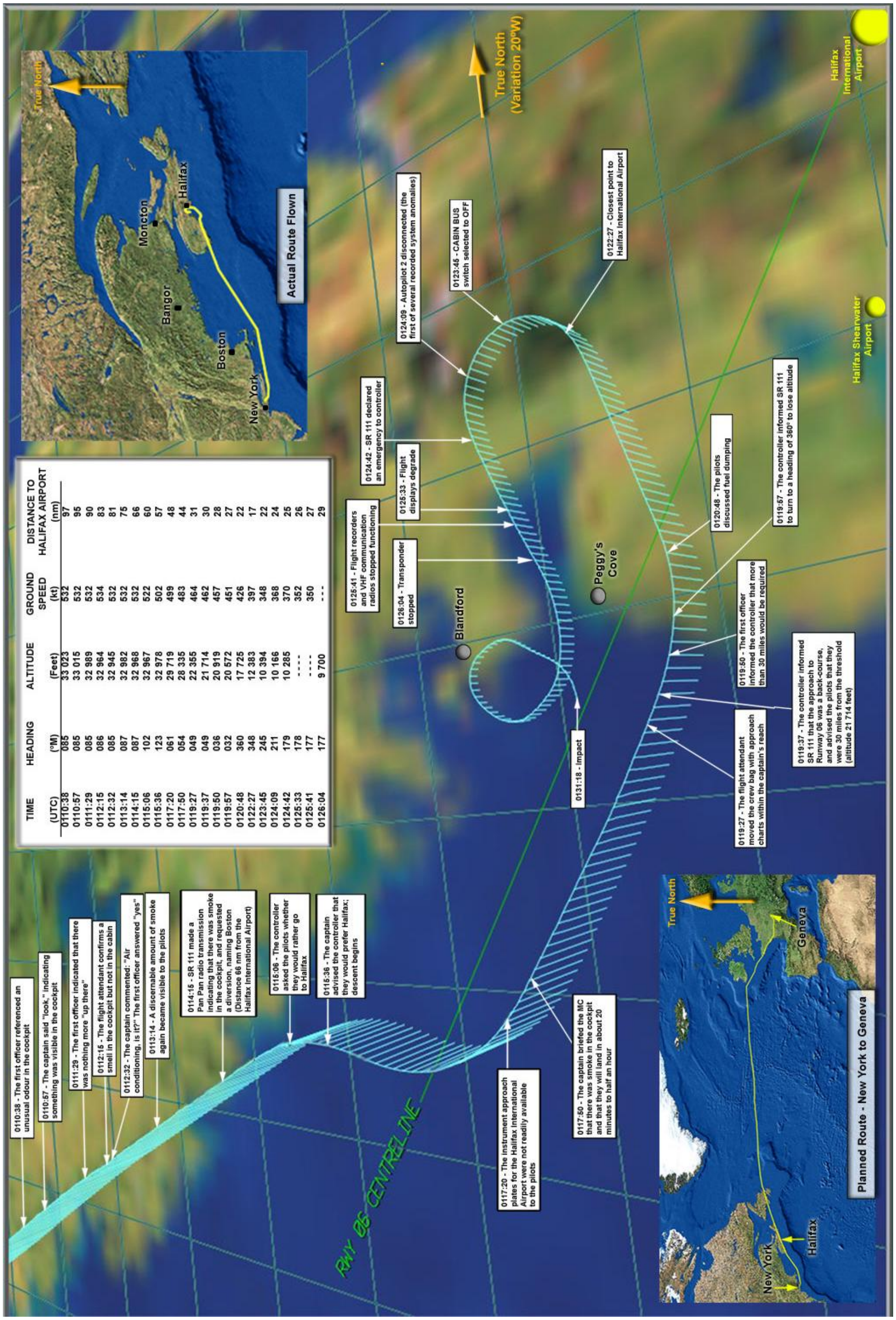


Figure 2: Flight Path of Swissair Flight 111 [1].

to lose altitude, the fire spread and disabled a number of aircraft systems, which lead to the aircraft crashing into the water. It is possible that if the pilots had initially recognized Halifax as the most appropriate landing location and been able to quickly devise a trajectory to land there, then the plane may have been able to land before the fire disabled critical systems.

As this example illustrates, once an emergency has occurred, the selection of a landing site and a safe trajectory to that site are of critical importance. The pilot's high workload and limited computational capacity establish the need to provide automated assistance. Additionally, this is a task which pilots are very infrequently called upon to complete. However, the highly complex nature of the selection and planning tasks, as well as the uniqueness of each emergency, makes automation difficult. The input and oversight of a human operator is required.

1.2 Background

The responsibility for the safe completion of every flight rests with the pilot-in-command. In an emergency situation, this can be a very challenging duty. The pilots must monitor the aircraft systems, detect and resolve any failures, control an aircraft with possibly degraded performance, and coordinate with the cabin crew, airline dispatchers, and air traffic control. In addition to these tasks, the pilots must also plan and execute a trajectory that will result in a safe landing. These tasks are made even more difficult by the circumstances. For example, the pilots may feel a sense of physical danger, or the cabin environment may be a distraction due to smoke, heat or noise. Additionally, the aircraft's performance may be affected, resulting in degraded or inadequate handling qualities. This limits the relevance of past experience to the current situation.

In order to understand some of the difficulties that these circumstances present, a number of cognitive engineering models are reviewed. One of the most important factors is stress, which can have negative effects on the decision maker's cognition. The Cognitive Control Model describes how the context of the emergency dictates the way in which the planning task is handled [12]. This model describes the degree of control a person has as dependent on the context of the situation. The degree of control is determined in large part by the amount of subjectively available time and the familiarity of the situation [13]. Subjectively available time refers to the amount of time that a person perceives that he or she has available to take action. The amount of time perceived may depend on the objective amount of available time, the predicted changes in the system, the person's level of arousal and other factors. The degree of control is discretized into four control modes: scrambled, opportunistic, tactical, and strategic. The relationship between the amount of subjectively available time, familiarity of the situation and the control modes is shown in Figure 3.

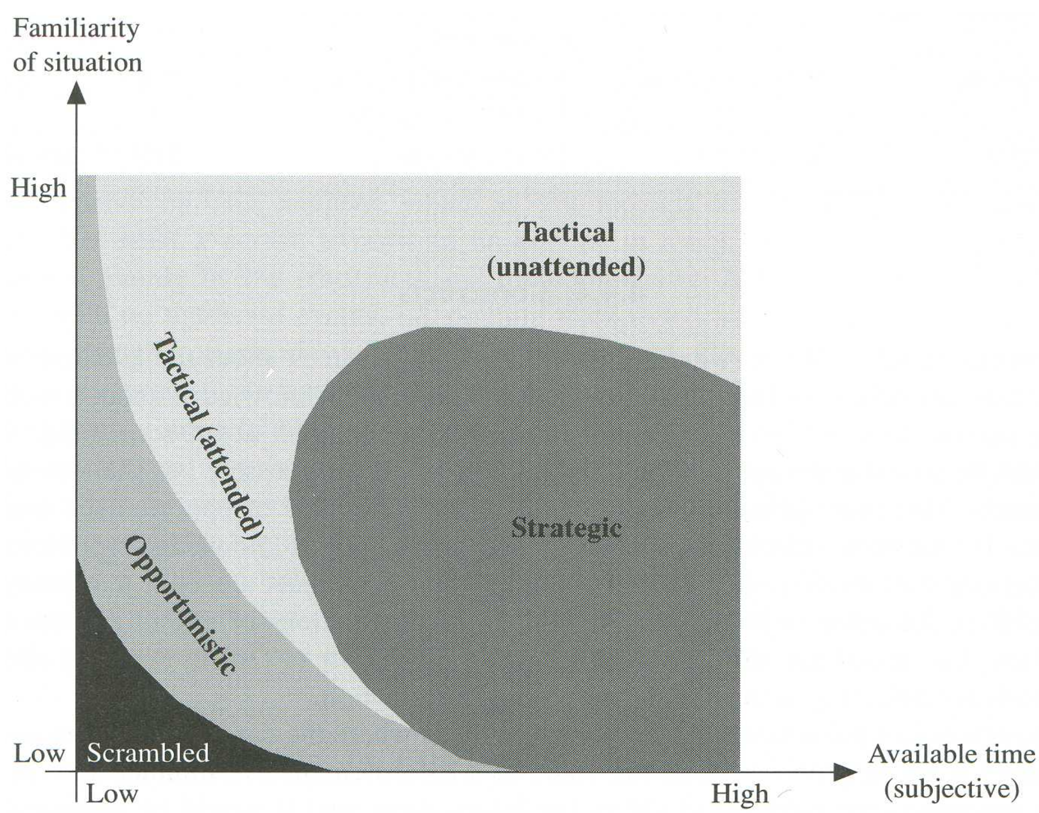


Figure 3: Relation Between Subjectively Available Time, Familiarity and Control Mode [13].

The simplest and most dangerous mode is the scrambled mode, which generally represents a person in a state of panic. When a pilot is in this mode, he or she is not be able to focus even on one goal of flying the aircraft. When a pilot has adequate subjectively available time, he or she may be described by the tactical mode. In this mode the pilot has a greater sense of control. The pilot is more likely to develop a plan or modify an existing plan in order to fit the current situation. The resulting plan may take into account the potential effects of candidate actions. This mode corresponds to "normal" performance. During an emergency situation, a pilot will likely be somewhere in between the scrambled and tactical modes, described by the opportunistic mode. In this mode, pilots may use the plans and procedures available; however, they may not be used correctly or effectively.

The amount of subjectively available time that a pilot perceives may be influenced by a number of factors. The phase of flight during which the emergency occurs, the type of emergency, the number of actions the pilot is required to complete, the availability of resources and the stress level all act to reduce the amount of subjectively available time. The stress may be physical, such as smoke in the cabin or loud noises, or it may be psychological, such as the fear of impending danger.

These stresses impact the manner in which the pilot makes decisions. While the pilot may be able to quickly develop a plan of action based on experience and intuition, stress can lead him or her to fixate on one solution, and fail to compare alternatives [4, 21]. Additionally, the pilot may simply increase the speed with which he or she processes information, potentially leading to errors. The pilot may also reduce the amount of information that is sought and processed, known as filtration [23, 24].

Stress may also lead the pilot to rely too heavily on an automated tool. He or she may assume that the plan generated by automation is best, without verifying its feasibility or exploring other options [25]. Also, the pilot may seek only information which confirms the automation-generated solution as the best, while discounting other information (confirmation bias). Alternatively, rather than simply discount conflicting information, the pilot may attempt to rationalize and mentally force all available information to fit the automation-generated solution (assimilation bias)[7]. These stress-related factors can cause pilots to make poor decisions, despite the fact that they would be able to make acceptable decisions under normal circumstances. These poor decisions may cause incidents to become accidents.

In addition to the stress of the circumstances, the complex nature of the decision making task is also important. The Naturalistic Decision Making framework is often used to describe how experts make complex decisions. Zsambok [30] describes NDM as, "the way people use their experience to make decisions in field settings." While experts are often able to make excellent decisions based on experience and intuition, many of the aforementioned effects of stress can negatively impact the quality of the decision.

An alternate model of decision making is the rational, analytic model. This model describes the decision maker proceeding through a set of steps to reach a decision. These steps generally include generating alternatives, envisioning the consequences, evaluating the alternatives against a set of criteria and choosing the best plan [18]. In general, a rational decision process may be helpful in determining the safest path to land, however, it may not be a reasonable undertaking during an in-flight emergency. A hybrid of the two decision models, taking the best of both worlds, may allow the decision maker to reach the best result. In a hybrid model, the rational decision process can compensate for some of the weaknesses of NDM. For instance, the rational decision process generates a number of alternatives, which alleviates the tendency to fixate on a single solution. By automating the generation and some of the evaluation of alternatives, the process can be done quickly. It should be noted that, as Peter Simpson [28] warns, "A decision aiding system should not become a decision making system, and it should never simply dictate decision courses to the operator."(p.18)

However, by capitalizing on the automation's fast lookup and simulation abilities, and the human's pattern recognition and intuition capabilities, the two may be combined to more reliably make sound decisions.

1.3 Related Work

In order for an automated planning aid to be most useful, there are at least two primary tasks which it must be able to accomplish; first, it must be able to accurately predict the most appropriate alternative landing site, as well as the most efficient trajectory to land at that site, and second, it must provide an interface with the pilot. The completion of the first task requires that the aid determine the overall feasibility of a trajectory, avoiding the case of a controlled flight into terrain because of not reaching the destination or overshooting it. A feasible trajectory must also avoid obstacles, which may be static, such as a mountain, or dynamic, such as a severe weather system. The consideration of static obstacles avoids controlled flight into terrain, whereas the consideration of dynamic obstacles avoids accidents as a result of flight into convective weather. The determination of such a trajectory must be made by taking into account the aircraft's possibly abnormal aerodynamics. Finally, the trajectory must minimize time to land, which is important in many cases, such as the aforementioned Swissair flight. The aid must also provide an interface with the pilot through which information is shared in both directions. Most research to date has primarily focused on one or the other of these tasks.

The landing site selection task has been suggested as a candidate for automation. Atkins, Portillo and Strube [3] have developed a method to complete this task. First, the footprint containing all feasible landing sites is calculated. Then the list is prioritized according to a number of weighted criteria, such as runway length, airport facilities available, etc. In their research, the authors chose example values for the criteria weights, but acknowledged that the criteria weights would ultimately be based on expert knowledge and would vary by emergency type.

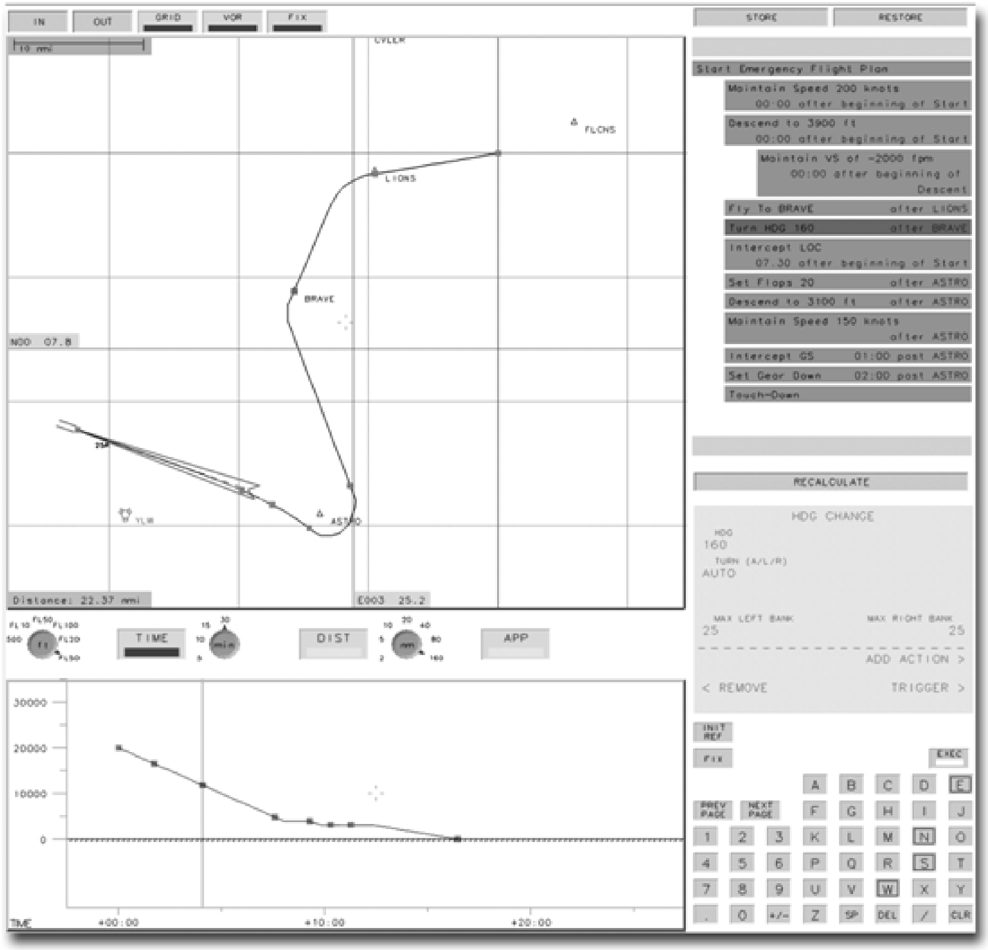


Figure 4: Emergency Flight Planner Interface [6].

The need for the pilot and the automated planning aid to interact with each other has also been investigated. The Emergency Flight Planner (EFP) by Chen and Pritchett [6] has been proposed as a prototype interface between the pilot and an aid. The EFP, shown in Figure 4, allows the pilot to enter a plan. The ensuing trajectory is then predicted and evaluated. The EFP also provides an additional mode in which the pilot is presented with a preloaded trajectory, which can then be accepted, modified, or deleted. The results of testing with the EFP emphasized that generated plans must incorporate the structure and objectives used by pilots in order to be effective.

Research reported by Layton, Smith and McCoy [22] in their study of a cooperative problem-solving system for en-route flight planning investigated three possible system modes of pilot-system interaction. In that study, pilots and air traffic controllers were both used as subjects. The study evaluated three possible modes. The first mode was a sketching-only system, in which a plan devised by the subject was evaluated by the system and feedback was provided. The second was a sketching system with the additional capability for the user to specify constraints on the plan and allow the system to propose a solution which matched those constraints. In the third mode, the system proposed a plan based on system-specified constraints. The results showed that in the second and third mode, users explored more possible options; however they were also biased toward the system-generated alternative. The same study also highlighted the fact that the use of a fully automated aid could be detrimental if it performs suboptimally.

The previous results show that, to increase the usefulness of an automated planning aid, the process by which pilots select an alternative landing location and plan a path to it must be better understood. In addition, it must be understood how the pilot can best be aided by such a tool. It is expected that an aid that accepts and provides information in a manner that is most consonant with the pilot's mental process will be most effective.

CHAPTER II

PILOT FEEDBACK FOR AN AUTOMATED PLANNING AID SYSTEM IN THE COCKPIT

In order to inform the design of an Automated Path Planning Aid, a survey was conducted by in order to better understand the tasks and priorities of pilots during an emergency situation [29]. The first section of the survey was intended to elicit information about the primary factors that pilots consider in the process of planning a landing trajectory. This involves choosing the most appropriate destination at which to land, and then determining a trajectory to reach the ground safely. The trajectory planning task also requires attention to certain en route considerations, such as severe weather and hazardous terrain.

The first section of the survey was structured to cover two general types of emergencies: 1) a performance altering scenario, in which the aircraft's performance was non-nominal, and 2) a non-performance altering scenario, in which the aircraft's performance was normal, but an immediate landing was necessary. For the non-performance altering scenario, the participants were presented with the following information:

> You are the captain of a civil transport aircraft. A fire has been detected in the cargo hold. The appropriate checklists have been performed, but the fire has not been completely extinguished. The first officer is controlling the aircraft, allowing you to plan a course of action.

For the performance altering scenario, the participants were presented with the following information:

> You are the captain of a civil transport aircraft. The right engine of your twin-engine aircraft has failed. The appropriate checklists have been performed. The first officer is controlling the aircraft, allowing you to plan a course of action.

The same set of questions was used in each of the two scenarios. The performance altering scenario also included an additional question, which addressed how the pilot would assess the feasibility of a trajectory given the aircraft's degraded performance.

The second section introduced the concept of an Automated Planning Aid (APA). The questions built upon the performance altering scenario from the first section, with the following additional information:

> Now an Automated Planning Aid (APA) is available to assist you with the selection (and perhaps execution) of a suitable plan of action.

This section was intended to obtain information about how the participants might use an APA. In particular, how participants prefer to convey information to the APA and furthermore how they prefer to review the information provided by the APA. Finally, this section presented questions meant to ascertain the amount of confidence that participants would have in the APA.

The third section was designed to collect further information about how an APA might be used. The participants were presented with the following information:

> Consider an emergency scenario which is unforeseen (i.e., you have not received any pertinent training). The aircraft's performance is now altered and/or degraded in some way. You are the captain, and the first officer is controlling the aircraft, allowing you to plan a course of action. In this scenario you do not have an Automated Planning Aid (APA) available to assist you.

This scenario was included because it provides some insight into how the participants will make a plan in a situation where they cannot rely on any prior training or procedures to guide them through the process.

The final section included general questions about the participants' opinions of the proposed APA concept. These questions asked about the scenarios under which the participants would be more willing or less willing to seek help from an APA and how the participants would like the plan to be updated. This section also included biographical questions in order to determine the demographic make up of the participants.

2.1 Methodology and Participants' Profile

The survey was conducted using an on-line service. The service was used to create, format, and monitor the survey. It was also used to host the survey and collect responses from the participants. In order to generate responses from several airline pilots, a link to the online survey was distributed via email to a number of pilots. In addition to the email, a link was posted on the airline pilot association's message board requesting participation. Responses were collected over the course of approximately six weeks between August and September 2008.

Responses were received from twenty-one participants. One of the respondents declined to include biographical information, however all twenty-one responses were used in the results. The demographic analysis therefore only includes twenty respondents. All participants held the position of either captain or first officer and had been in their current position for an average of 9.5 years. Eighty-five percent were flying a Boeing aircraft (737, 757, 767, or 777) at the time of the survey. All pilots had at least 6,500 flight hours with an average of 12,979 flight hours.

2.2 Survey Results

Due to the small number of responses, slight variations in the number of responses for a given option were neglected as statistically insignificant. However, each question also included an open-ended option where the participants were free to provide more information. These responses often provided additional valuable insights into the participants' thoughts that could not be captured by the multiple choice responses provided.

The first section of the survey included a set of questions about how pilots currently make decisions in an emergency situation. The question set was first given for a non-performance altering scenario, and then repeated for a performance altering scenario, as described previously. Participants were asked to indicate the priority (high, medium, low, or not a consideration) associated with a number of factors when choosing the airport to which the respondent would divert. Under the non-performance altering scenario, the most important factor indicated was the proximity of the airport in terms of time. Weather conditions at the airport, the length of the runway and the proximity of the airport in terms of distance were also given relatively high priority. These results can be seen in Figure 5. Under the performance altering scenario, the results were largely the same; however the importance of proximity in terms of time was not differentiable from that of other factors, as seen in Figure 6. One free response comment for the non-performance altering case also indicated that runway lighting and the availability of navigational aids were additional important considerations.

A similar question was posed for each scenario in which participants were asked to indicate the priority (high, medium, low or not a consideration) associated with a number of factors when planning a safe path. For both scenarios, en route weather and the avoidance of hazardous terrain were given the same priority, and low priority was given to traffic routes. This is not surprising because once an emergency is

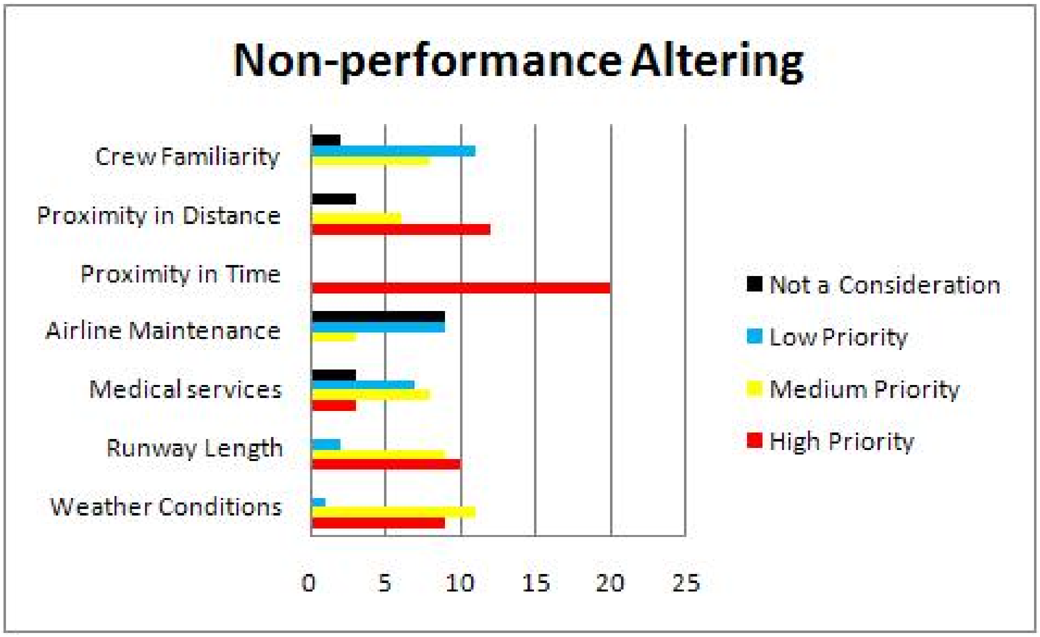


Figure 5: Prioritization of Landing Site Criteria for Non-Performance Altering Case.

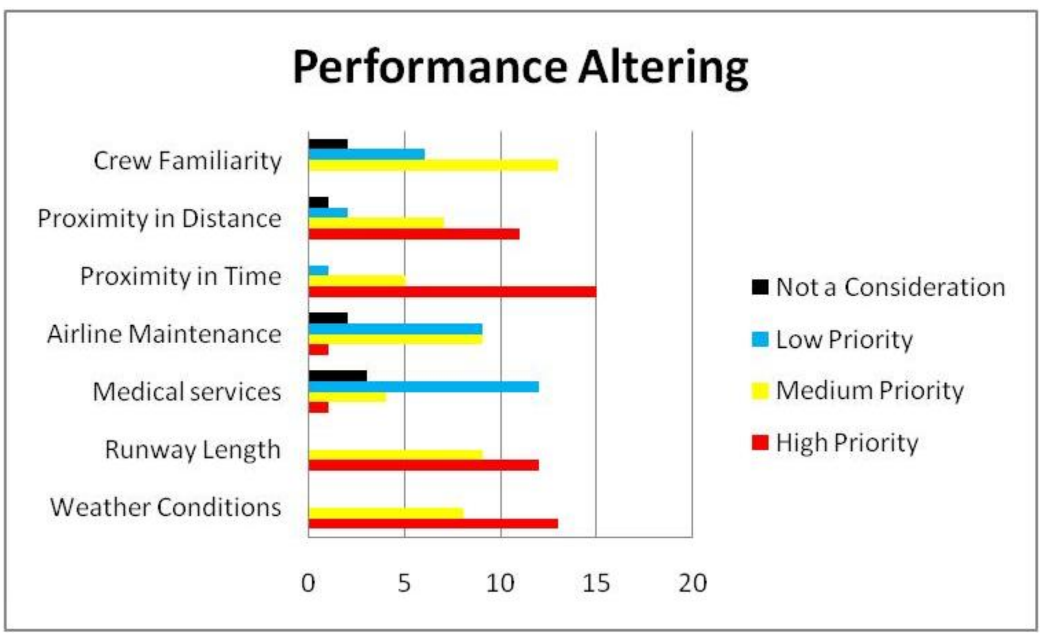


Figure 6: Prioritization of Landing Site Criteria for Performance Altering Case.

declared, the pilot need not comply with ordinary routes and approach procedures.

The performance altering scenario differs from the non-performance altering scenario in that the pilot's experience and knowledge of the aircraft may have limited applicability to the current situation. For this reason, the pilot's "first instinct" may be the best plan given normal performance, but may not be feasible given the degraded capabilities of the aircraft. Participants said that they were most likely to judge the feasibility of a maneuver by running the scenario mentally or seeking help from the dispatcher. Many pilots would also consult the performance manuals. These responses indicate that an automated planning aid may be particularly helpful in situations where the aircraft's performance is not normal.

Participants were then asked to consider whether they would completely determine a plan before taking any action. The alternative would be for the pilot to alter the current course immediately based upon his "first instinct." Many of the respondents took advantage of the open-ended option to describe some other considerations that affect how they would proceed. For instance, when flying in a mountainous area such as Quito, Ecuador, planning ahead is essential. Others indicated that they would first coordinate with Air Traffic Control (ATC) before taking action or completing a plan of action. Those who would first take some action responded that they would all turn toward the nearest airport. None would begin descending. For the non-performance altering scenario, most of the pilots reported that they would change the current course immediately. However, in the performance altering scenario, most would develop a plan first before altering the current course. This may indicate that pilots are more comfortable taking immediate action in a more familiar situation, as opposed to a novel scenario. For instance, pilots have trained for an engine failure scenario and would have a relatively good idea of how to control the aircraft. However, in the event of a control surface malfunction, they would not be as familiar with the aircraft's post-failure performance and may be less likely to take immediate action before planning.

The final question of this section addressed the parts of a potential plan with which the participants would be most careful (i.e., provide more specific attention, add more detail, etc.). The pilots were asked to indicate the priority (high, medium, low, not a consideration) associated with portions of the plan. The results are shown in Figures 7 and 8. In both scenarios, pilots reported that the highest priority is around severe weather and hazardous terrain. Medium priority was given to the area around the aircraft's current location. In the case of the performance altering scenario, medium priority was also given to the area around the destination.

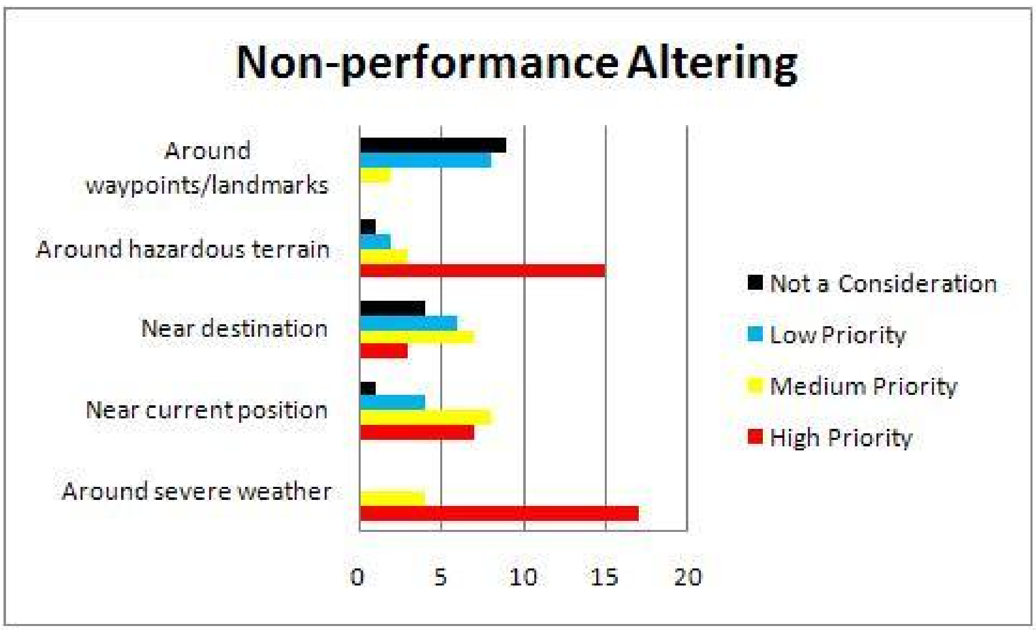


Figure 7: Parts of the Plan Pilots Consider with Detail for Non-Performance Altering Case.

The first questions of the section dealing with an Automated Planning Aid (APA) address the interface between the pilot and the APA. Pilots described a number of inputs as either highly preferable, somewhat preferable or not preferable. Pilots will need to be able to provide priority information to the APA in order for it to be aware of the current situation. Participants indicated that they would prefer to accomplish this either through the Flight Management System (FMS) pages, through a separate

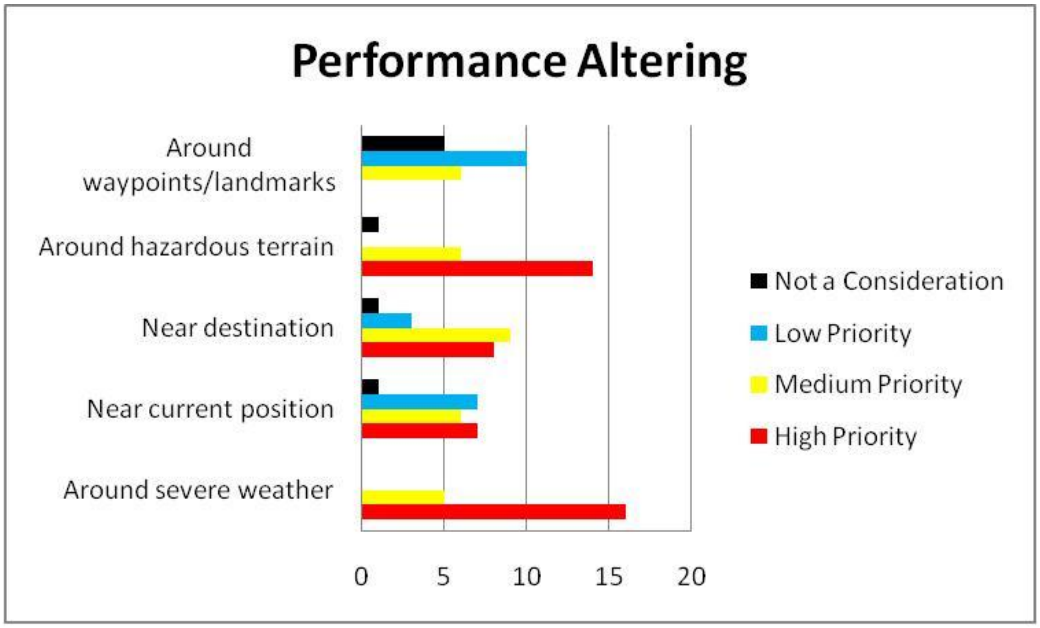


Figure 8: Parts of the Plan Pilots Consider with Detail for Performance Altering Case.

dedicated interface, or through a data link from the airline operations center. When the APA has developed a plan, the pilot must be able to review this plan. The respondents indicated that they would prefer to review the plan as a set of automatically generated FMS entries. The pilots would also favor a horizontal graphical representation of the proposed trajectory, possibly accompanied by a vertical profile of the proposed trajectory.

In order to effectively evaluate the proposed plan provided by the APA, pilots will have certain metrics in mind which will be used to make the evaluation. Participants were asked to indicate how important certain metrics are when evaluating the plan (highly important, somewhat important or not important). The most important metric was the cumulative time/distance and fuel information. A comparison against alternatives was also considered an important metric.

Responses showed that the pilots will not unconditionally follow the plan generated by the APA, especially if the plan if different from their "first instinct." Most participants said that they would follow an APA-generated plan that was different

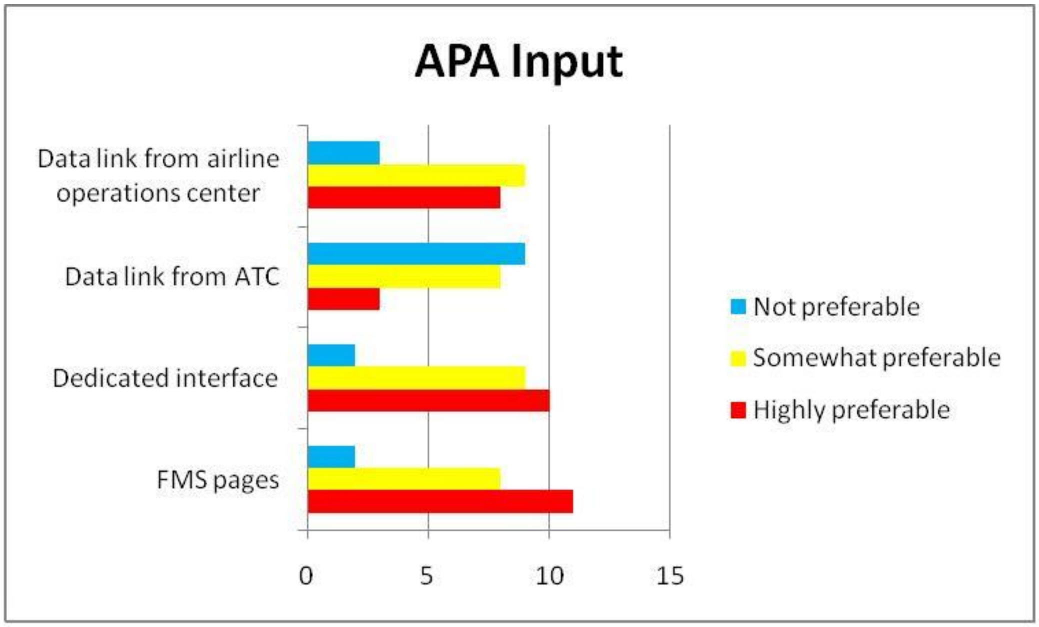


Figure 9: Preference for APA Method of Input.

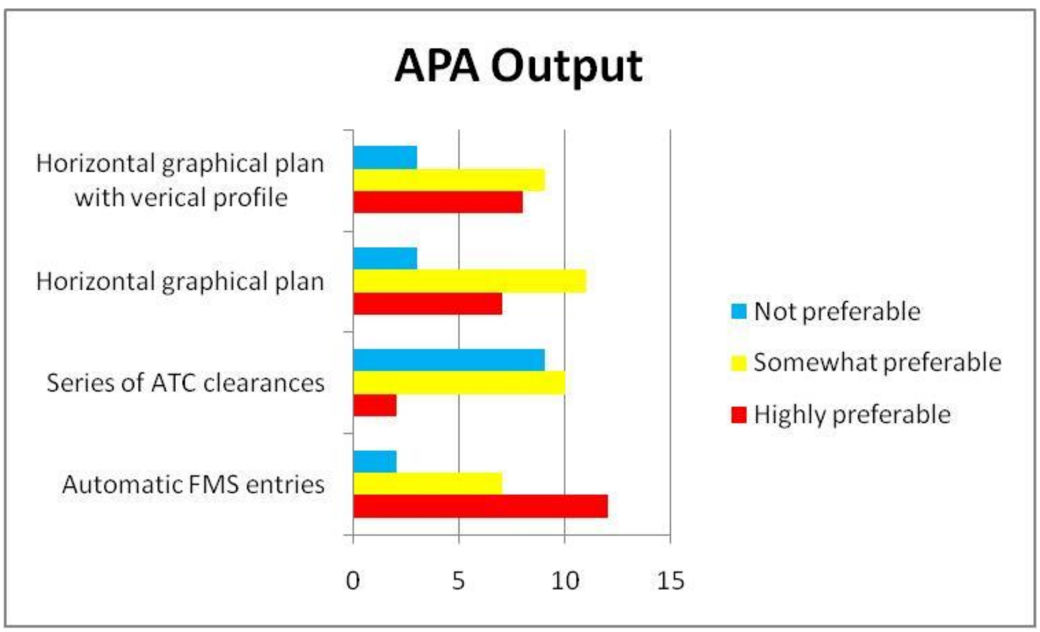


Figure 10: Preference for APA Method of Output.

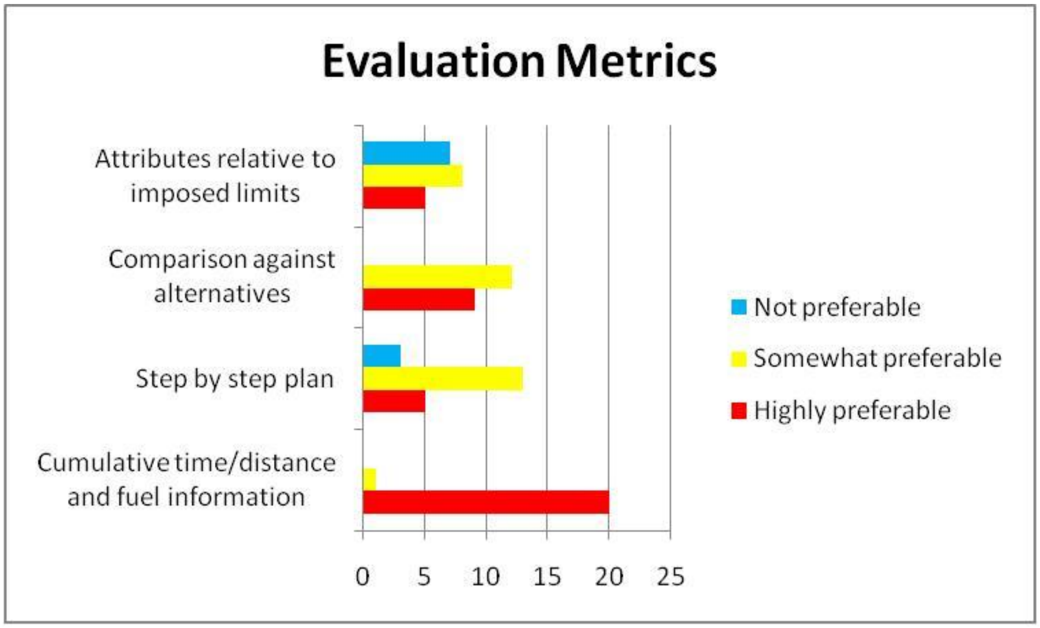


Figure 11: Metrics Used to Evaluate a Plan.

from their own when the APA plan required a significantly shorter amount of time to execute. Some pilots also said that they would follow the APA plan if it remained well within the flight envelope limitations and encountered significantly less severe weather. Nearly all of the pilots indicated that they would only use the APA-generated plan as an aid; that is, they would take it into account while re-evaluating their own plans of action, but would neither completely accept nor reject an APA-generated plan.

In the final section, the participants were given an unforeseen emergency, as described previously. As in the performance altering scenario, the majority of respondents would completely develop a plan before altering their course. Many participants again took advantage of the open-ended response option to indicate that the primary factor in deciding whether to take immediate action would be the urgency of the situation. Of those who would immediately alter the current course, most would turn toward the nearest airport.

In a situation in which the pilot has accepted the plan, but has deviated from it, a new, more efficient, plan may be calculated as a result of the deviation. Pilots were asked to choose between a number of conditions under which they would like for the APA to provide a new plan. The results are shown in Figure 12. Most pilots indicated that they would prefer to only receive a new plan from the APA when they asked for one.

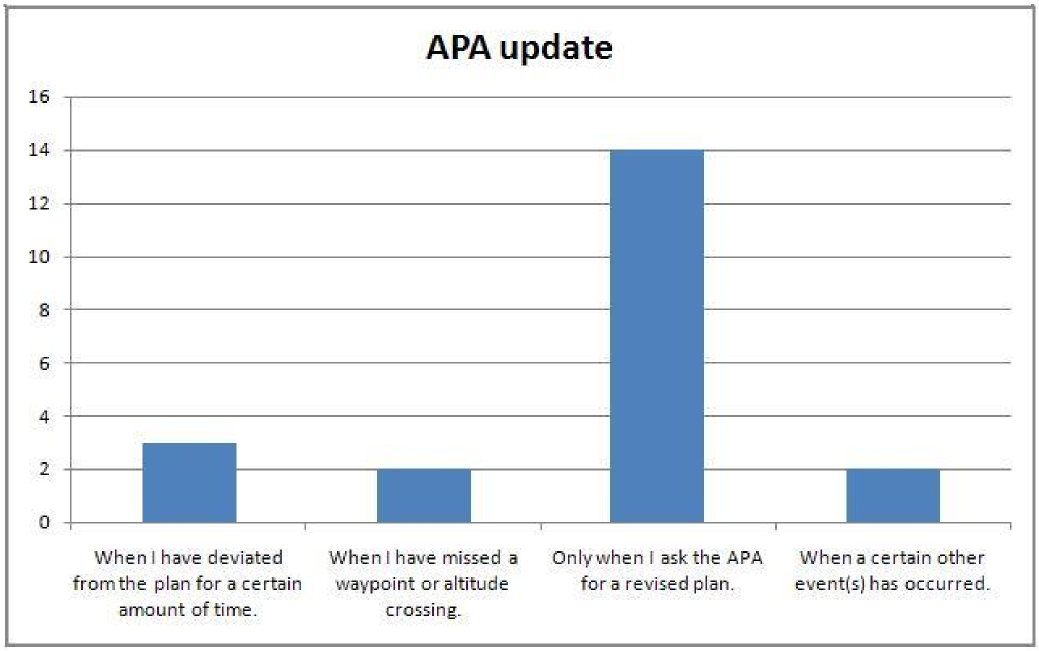


Figure 12: Conditions Under Which the APA Should Provide an Updated Plan.

Not surprisingly, the majority of pilots reported that the situation in which they would be most likely to rely heavily on the APA is one for which they have not had any prior training or experience, as shown in Figure 13. Some indicated that they would not rely on the APA in either a familiar or unfamiliar situation, while others said that they would rely on the APA in both situations. This may be due to the lack of clarity with regard to reliance. It seems, based on these comments, that many of the pilots would use the APA as an aid, but they would be hesitant to simply follow its plan without some verification of their own.

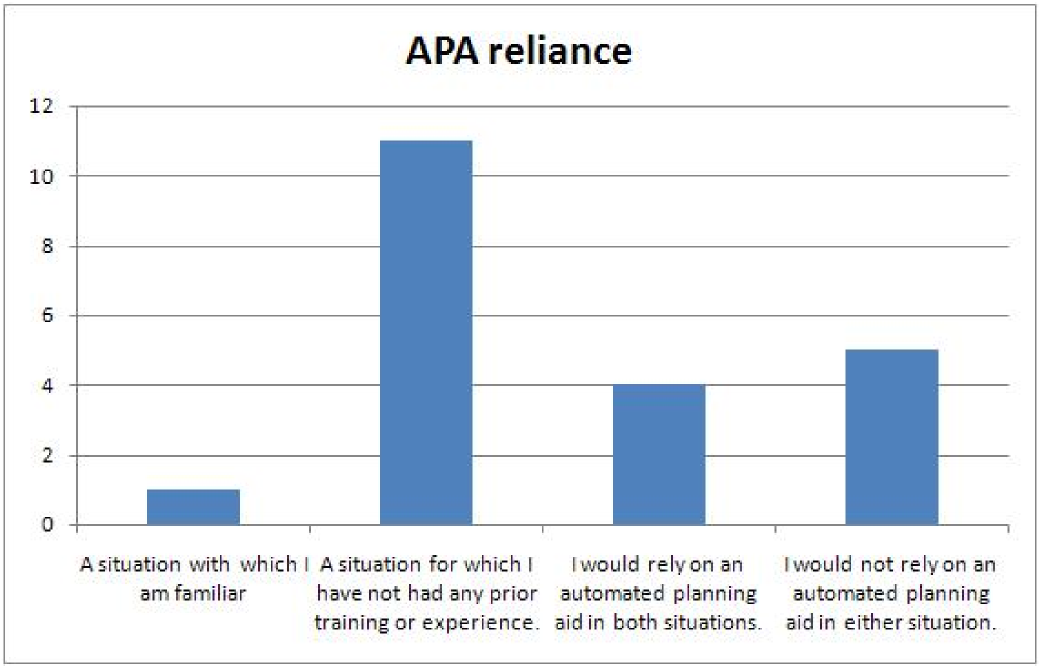


Figure 13: Situations in Which an APA Would Be Most Heavily Relied Upon.

2.3 Analysis of Responses

The multiple choice nature of the responses made the mathematical results simpler to discern; however, the more enlightening portion of the responses were the open-ended options, which allowed the participants to include their thoughts on each of the questions provided. These responses provided insights into the pilots' expectations about how an APA should function and how it could be most useful.

As mentioned previously, the pilot's workload is very high during an emergency. A number of comments indicated that pilots are wary of any factor(s) that would necessitate more work during a stressful time. This is reinforced by the respondents' preference for interacting with the APA through the FMS, a device they use for other purposes, and with which they are familiar. Most participants preferred to review the plan as automatically-generated FMS entries, which would allow the autopilot to follow the APA path with very little additional work required by the pilot. Respondents emphasized that the APA should be an efficient source of information which is currently disparate, if available at all. One pilot commented that, "it should offer information, but not demand any acceptance or response."

In general, commercial pilots have a good working knowledge of the areas in which they normally fly. For this reason, the pilots' "first instinct" is often very good. The results concur with this conclusion, as evidenced by the fact that pilots were more likely to alter the current course without a complete development of a plan in the non-performance altering case and much less so in the unforeseen case. A number of comments revealed that in some cases pilots simply need a tool to validate their plans and point out to them any options that they may have missed. A tool such as this may have been helpful in the case of Swissair flight 111, which initially turned toward Boston when an emergency was declared despite being closer to Halifax, Nova Scotia [1].

When determining the best landing site, as well as the best path to that site, a number of pilots found the list of factors provided to be insufficient. Certainly, the factors mentioned are important, however, some comments emphasized the reliance on outside sources. Once an emergency has been declared, pilots work very closely with Air Traffic Control (ATC) to receive their input to determine the most appropriate path. Pilots' comments also indicated that they will seek advice from the airline dispatcher in order to determine the most appropriate landing site. The scenario for which the pilot has not had any training or prior experience garnered a number of additional comments. These emphasized the interactive nature of the planning task by pointing out that the process must include ATC, airline dispatcher, other crew members, and possibly the manufacturer.

Many comments addressed the role of the APA in the planning process. In addition to keeping the workload as low as possible, many pilots do not want to view an APA-generated path as a directive. Rather, they prefer to view it simply as one input into the process of developing their own plan. This supports the result seen previously in the literature, in which nearly all participants said that they would take the automated plan into account while re-evaluating their own plan. The comments emphasized that the path planning task is complicated and that the automated tool may not have the ability to gather a complete understanding of the situation at hand. These comments imply that pilots would like to have an APA that makes critical information easy to access in a timely manner, but which does not dictate actions for the pilot to follow.

Due to the uniqueness of each emergency, one respondent proposed an approach which would likely be supported by other pilots. The pilot said, "You should be able to manipulate individual variables and compare solutions." The priority of certain criteria may change, depending on the emergency, and pilots need to be able to indicate this to the APA. For instance, airport fire and rescue services are more important in a fire emergency. Runway length may be more important in a flap or landing gear malfunction. The recognition that there are a number of variables which must be taken into account was echoed by a number of comments. Also, the ability to compare alternatives was considered an important metric.

A fundamental requirement of an APA is that the pilots must trust it and must be willing to use it. One pilot addressed this issue by saying that, "I must know how the automated plan is generated to be able to trust its output. Once I have confidence in the APA planning process, I would be more likely to trust its output, particularly in a time-critical emergency situation." This sentiment would surely be echoed by other pilots who will be hesitant to trust any tool with which they disagree. It is these disagreements that provide the usefulness for an APA; if the generated plan always agrees with the plan the pilot has in mind, then the tool has provided only a very limited service. However, if the pilot and the tool disagree, the tool must be able to demonstrate to the pilot that his plan can be improved upon. It must also be ensured that the pilot, working with the APA, does in fact generate a better plan than the pilot could on his own. This must be done without causing the pilot to completely rely on the system through biasing or over-reliance [27].

2.3.1 Criteria Weighting

In order for an Automated Planning Aid to develop a recommended path, the automation must be able to develop a prioritization among the possible landing sites. Such a prioritization may be based on the minimization of a utility function, such as Equation 1.

$$U = C_1 * \frac{t}{t_{max}} + C_2 * \frac{d}{d_{max}} + C_3 * \frac{r_{l,max} - r_l}{r_{l,max}} + C_4 * w_{wx} + C_5 * w_{cf} + C_6 * w_{med} + C_7 * w_{rep} \quad (1)$$

In this equation, the seven parameters are: time required to land t, distance from the current location d, runway length r_l, weather conditions w_{wx}, crew familiarity with the landing site w_{cf}, medical services available w_{med}, and airline maintenance and repairs available w_{rep}. The time t and distance d are nondimensionalized by their maximum possible values given the aircrafts performance t_{max} and d_{max}. The availability of medical services and airline maintenance are static attributes of each airport which could be encoded on a scale from zero to one. The value for the crew familiarity factor could also be assigned before a flight by the pilots. The weather factor must be determined in real-time, based on the probability and severity of adverse weather conditions.

This leaves the determination of the criteria weights C_i. In the survey, each of the landing site criteria were assigned a priority; high, medium, low, or not a consideration. For this analysis, each of these options was assigned a value, three for 'high', two for 'medium', one for 'low' and zero for 'not a consideration.' The responses were summed based on the assigned values, giving a total for each criteria. In order to normalize these values, each was divided by the sum of the total scores for all criteria, as in 2. The subscripts of *score* in Equation 2 refer to the score assigned by respondent r for criteria i. The resulting values for the criteria weights are shown in Table 1.

Table 1: Resulting Criteria Weights Based on Survey Results

	Non-performance Altering	Performance Altering
Proximity in Time	0.21	0.19
Proximity in Distance	0.17	0.16
Runway Length	0.18	0.18
Weather Conditions	0.18	0.18
Crew Familiarity	0.10	0.11
Medical services	0.11	0.08
Airline Maintenance	0.05	0.10

$$C_i = \frac{\sum\limits_{r=responses} score_{r,i}}{\sum\limits_{j=criteria} \sum\limits_{r=responses} score_{r,j}} \tag{2}$$

These weights were used to calculate scores for the landing sites in eight scenarios. These were the landing sites used in the experiment described in Chapter 4. For the experiment, the landing sites were ranked by an expert (see Section 3.2.1.2). The expert's ranking was compared to the ranking resulting from the implementation of these criteria weights. This weighting system performed reasonably well, selecting the same "best" landing site in six of the eight scenarios. To assess the overall match between the two rankings, the distance between the two rankings that each site was given was calculated (this is the absolute value of the difference). The overall average distance was 1.27 ranking positions. The rankings for two of the scenarios differed significantly, each with a total distance of 11 ranking positions between the six landing sites.

This is just one possible utility function. There are certainly other criteria that could be considered. For instance, one comment indicated that runway lighting and the availability of instrument landing system equipment may also be taken into consideration. This survey has provided one possible starting point for the relative weighting of these criteria. However, the determination of the most appropriate weighting and utility function must be investigated further before it can be considered for implementation.

2.4 Survey Summary

This survey investigated the pilots' tasks in the event of an in-flight emergency, namely the tasks of choosing a safe landing site, and developing a safe trajectory to reach that site. The survey has produced results that give some perspective into the methods and priorities pilots use to accomplish these tasks. Some insights into the manner in which an automated planning aid may be most effective have also been obtained.

During an airborne emergency, the need to land quickly is always of high priority. Therefore, the most important factor considered by the pilots when selecting an alternative landing site is proximity in terms of time. Additionally, the weather at the airport, the length of the runway and the distance from the current location are also important criteria. The most important en route factors are the avoidance of severe weather and hazardous terrain. However, each emergency presents a unique and complex scenario. When these criteria conflict with each other, the pilot must alter these priorities according to the situation at hand. It is the dynamic priorities based on situational awareness and experience that is difficult to capture and automate in a general sense.

A pilot's experience and expert judgment make him an effective decision maker in the face of emergencies. In a situation with which a pilot is familiar, such as a non-performance altering scenario, the pilot is more likely to make an accurate analysis and act quickly. Alternatively, in a situation with which he is unfamiliar, he will likely be more cautious and seek the assistance of an automated planning aid. In either case, an APA can be a valuable resource for the pilot to receive information and evaluate plans.

When the pilot is evaluating a plan, time is the highest priority. Therefore, the most important metric used is the cumulative time required to land. However, due to the complex nature of most emergency situations, tradeoffs should be allowable or even necessary. Consequently, the ability to compare a plan to other alternatives is also useful.

One of the most important aspects to be considered in an emergency situation is the high workload, time-critical, stressful nature of the situation. Accordingly, one significant feature of any proposed aid is that it should reduce workload, rather than increase it. The aid must provide useful information in a coherent manner, without burdening the pilot with requests. Similarly, pilots view the aid only as a tool, not as a directive. Pilots will use an automatically generated plan in conjunction with their own experience and intuition. Ultimately, the pilot has the final decision-making authority.

2.5 Design of an Automated Path Planning Aid

Based on these results, an initial design of an automated path planning aid was developed. The design of a new tool to be used in the cockpit is a very complex task. The amount of information and controls available in a modern cockpit is quite large. Also, the physical area in which they must be contained is rather limited. All of the systems' displays and controls must be contained in a small and coherent cockpit layout. With this in mind, no single part of the entire system should be designed on its own. It does not exist as a stand alone entity, but must work cooperatively to allow the pilots to complete all of their responsibilities.

An initial conceptual design was developed which captures the manner in which the tool supports the tasks, without constraining its implementation. This follows the User Environment Design described by Beyer and Holtzblatt [5]. The UED modeling technique highlights focus areas which "show the coherent places in the system that support doing an activity in the work."(p. 306) The model also shows how these focus areas are connected.

This design is comprised of the six focus areas as shown in Figure 14. The focus areas shown outside of the dotted box are outside the current scope. The first focus area is the place in which information is provided to the tool about the emergency. This information may be a simple indication of the urgency of the situation. For instance, in the case of very dire emergency the aid must know to consider any and all options. This information is used by the tool in order for the most appropriate results to be generated.

The next focus area is the place in which the user compiles a list of possible destinations and plans. The possible options may be displayed in a manner which indicates ranking according to some overall criteria. Details about each of the possible options is available. Databases about the surrounding airports and terrain help inform this list. This list automatically filters out infeasible solutions. The plans which are automatically generated take into account the aircraft's post-failure dynamics, as reported by the onboard systems.

The user may elect to generate a new plan or to modify a previously generated plan. The third focus area is where the user completes this task. The user may select a destination that was not included in the compiled list. The system generates a candidate path to that destination, which the user may alter. Based on all the solutions available, both user defined and automatically generated, the user then selects a candidate solution.

After choosing one of the possible plans to be the candidate, the user then moves to the evaluation focus area where the candidate plan can be evaluated. Only plans that the user has modified need to be evaluated by the system. The current wind and weather information, as well as the model of the aircraft's post-failure dynamics are used to determine the feasibility of the proposed path. While this portion of the system is critical, the validation of user-defined plans is beyond the scope of the

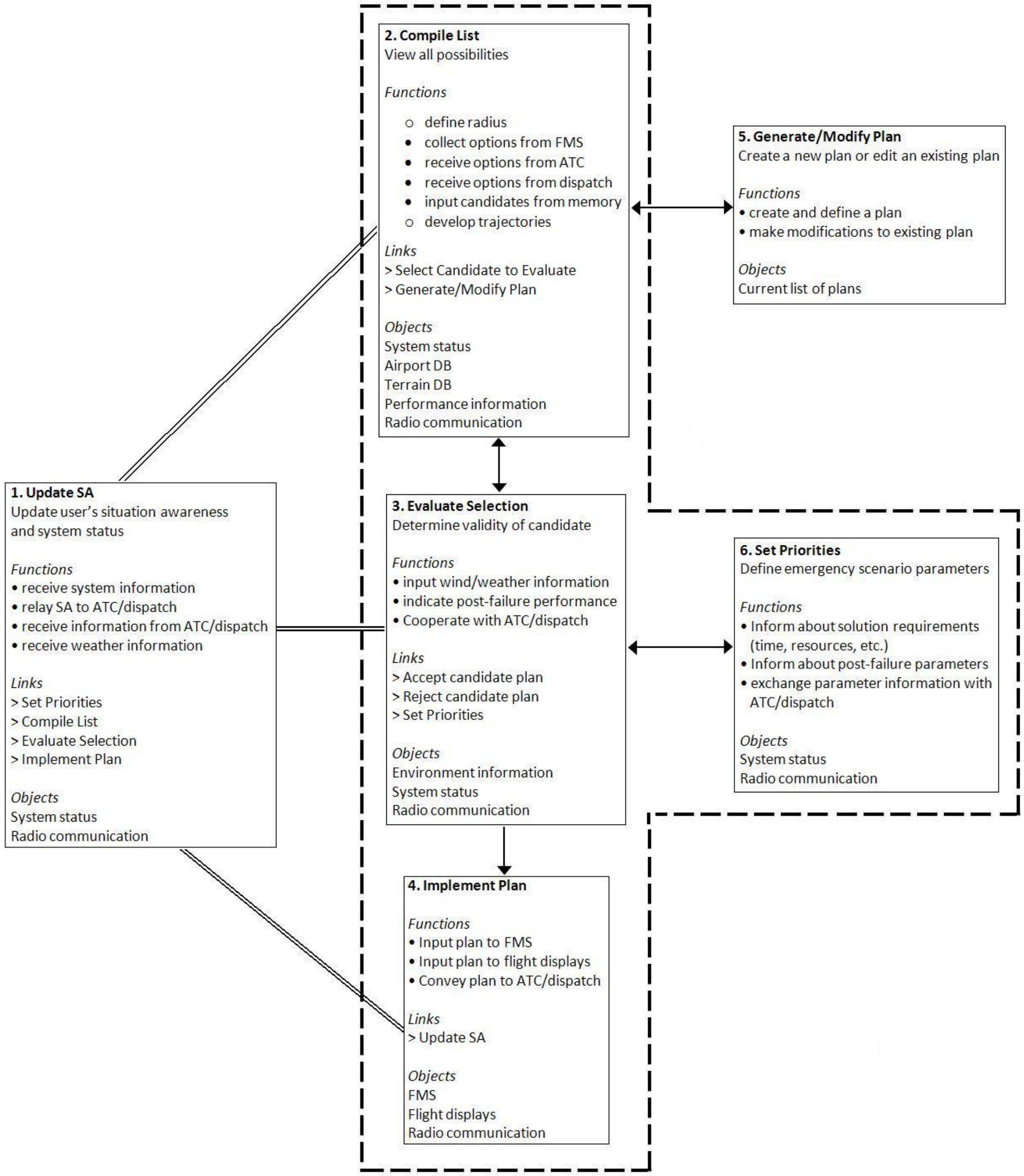


Figure 14: The User Environment Design.

current project. Once a plan has been deemed acceptable, the plan is implement by the system. It is distributed to other parts of the system such as the FMS or the primary flight display in order to be executed either by the autopilot, or manually by the pilot, using the Flight Director.

The user does not necessarily visit each focus area once and only once. The use of the tool in a specific situation is largely dependent on the user's mode of control. In an opportunistic mode, the user may only compile a list and implement a plan. Conversely, in a more strategic mode, the user may return to one or more focus areas as the situation evolves.

This initial design concept was used to develop a more detailed interface. The interface is intended to meet all of the design constraints highlighted in this conceptual design, while remaining as close as possible to the interface methods with which the pilots are already accustomed. The design will be discussed further in the following section.

CHAPTER III

TOOL DESIGN

3.1 Overview and Approach

It should first be noted that pilots are currently not without some form of automated path planning assistance. The modern Flight Management Systems (FMS) which are used on board most major transport aircraft include a couple pages to help the pilot with the task of deciding on a divert landing site. For example, the 'Alternate' page of the Boeing 777 FMS, shown in Figure 15, displays four possible alternates at a time [14]. They may be input from a list that the pilot creates before the flight, from a database, or entered manually. The estimated time of arrival (ETA) and predicted amount of remaining fuel are displayed. These four alternates are ordered by the ETA.

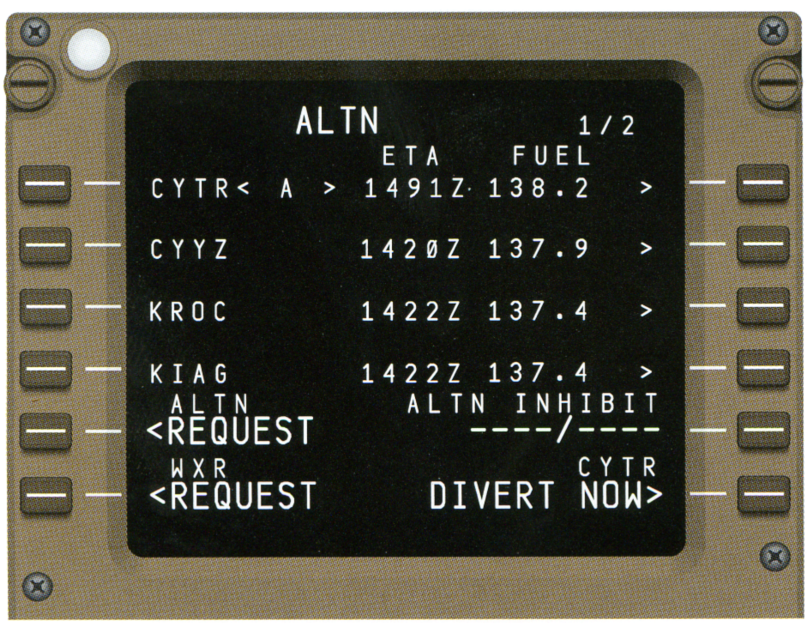


Figure 15: Example of Current Alternate Destinations Page [14].

While these pages are helpful, they leave much of the work up to the pilot. For instance, the nearest airport list only provides the landing sites at airports at which the aircraft is able to land normally, without taking into account the severity of the emergency. For instance, in the case of a severe emergency, the pilot may be willing to land at a runway that is not sufficiently long. Additionally, in the event of a performance altering emergency, such as a stuck elevator, the FMS cannot presently account for the post-failure flight dynamics of the aircraft. Thus, the plan that is generated may not be feasible given the aircraft's performance. The pilot may improve the plan by altering the waypoints used, however, this requires a good bit of time and work on the part of the pilot. In the case of an in-flight emergency, both time and cognitive resources may be limited due to the number of other tasks the pilots must address, which makes the current FMS solution impractical in highly time-critical emergencies.

In order to address these shortcomings of the present system, the following Automated Planning Aid is introduced. The APA first receives information about failures as they occur in the systems. This data is then used to determine the post-failure performance of the aircraft. These flight dynamics, combined with terrain and weather information, are used to compute time-optimal plans to reach a number of potential divert locations.

Each of the paths to the alternate landing sites is display graphically on the Navigation Display (ND) as well as textually in the Control Display Unit (CDU). Information about each landing site is collected from precompiled database information such as data about airports and terrain, as well as live weather information. This information is made available to the pilot through the CDU. Based on the information collected each site is given a score from 0 to 1, with 1 being best. From these scores, a cumulative score is calculated based on the system's weighting of the factors. The alternates are presented in descending order of overall score in the CDU.

3.2 Design Considerations

A number of the design decisions were made in accordance with the survey results and User Environment Design discussed in Section 2. Additionally, a successful design must be as closely integrated as possible with interfaces with which pilots are currently familiar. For these reasons, the APA was integrated into the existing Control Display Unit in a page which is based on the current ALTN page. The alternate routes are displayed on the existing Navigation Display per survey respondents preferences. Efforts were also made to ensure that the aid would not adversely contribute to the pilots' workload, but rather present relevant information in a coherent manner. With this in mind, focus area 6 of the User Environment Design was reduced to a single dial, which allows the pilot to quickly indicate the severity of the emergency by limiting the types of alternates which he or she is willing to consider. These levels are described in more detail in Section 3.4.3.

3.2.1 Scope

3.2.1.1 Calculation of Routes

In order for the ND to display these routes to alternate destinations, it must have some method of determining these routes. There are a number of possible approaches and a great deal of study has been done on this topic, such as [17, 20]. In the current software design, the approach taken was to calculate these alternate routes in real time using a Dubin's Path algorithm [9]. This simplified the trajectory generation routine without resorting to hard coded paths.

3.2.1.2 Criteria Weights

An algorithm for determining appropriate criteria weights based on type of emergency warrants a study of its own, and is not the focus of the current work. In order to avoid testing a specific criteria weight design, such as that derived from the survey results, cumulative scores were hard coded instead. A subject matter expert was consulted in order to determine appropriate cumulative scores for every site in each scenario. This expert had more than 20,000 hours of flight experience in over 20 years of service as a commercial pilot. The expert was provided all information available about each landing site and, unlike experiment subjects, was given an unlimited amount of time to consider each scenario thoroughly. These scores served to rank the landing sites and were the cumulative scores presented to pilots as a weighted combination of the criteria scores.

In addition, a design was considered in which the pilots would be able to alter the weights assigned to each criteria. This feature was omitted from the tested design in order to ensure that a difficult interface for manipulating weights did not impact the results of tested features. After the completion of all test runs and questions, pilots were shown one possible interface for manipulating weights and were asked to comment on the usefulness of this ability as well as the specific interface shown.

The specific interface design on which pilots commented is shown in Figure 16. In this design, the pilot could enter a weighting from 0 to 1 for each criteria under the constraint that the sum of weights must equal 1. After altering weights, such that the sum is 1, the pilot can update weights, at which time the cumulative score for each site is recalculated and the order in which the sites are presented in updated accordingly.

3.2.2 Technical Assumptions

A number of assumptions have also been made about the technological capabilities of the APA system. First, it was assumed that the APA has information about the status of pertinent aircraft systems, such as the engines and control surfaces. While this may be realistic in some circumstances, such as a complete engine failure, current systems may not be able to provide full information in other situations, such as a stuck aileron. It was also assumed, given knowledge of the failure, that the

Figure 16: Control Display Unit, UPDATE WEIGHTS Page.

post-failure flight characteristics can be accurately modeled. This model is used to determine the most time-optimal path that is feasible.

Assumptions have also been made about the knowledge that the system has of the environment. The system must have a complete and up to date database of all possible landing sites in the area of the flight. This database would include all airports, including small, uncontrolled airfields. This database would also include a terrain database, indicating the locations of large, clear areas where an aircraft could attempt to land in the most dire circumstances.

3.3 *Reconfigurable Flight Simulator*

The prototype which was used in testing was built using the Reconfigurable Flight Simulator (RFS), designed by Ippolito and Pritchett [16]. The simulator was developed in a modular manner allowing for the development of the necessary modules which could then be used in conjunction with the existing simulation engine, dynamics and control, and display modules. Each of the display modules used in this simulation is roughly based on the Boeing 777 type displays. The modules used in this experiment are described below.

3.4 *Displays*

3.4.1 Navigation Display

The Navigation Display (ND) is on the right of the screen in front of the participant and shown in Figure 17. This display provides a number of pieces of information. The display is track up, with the current plan shown in magenta. The graphical display of the information (primarily the routes to alternate destinations) allows the pilot to get an idea of the possibilities at a glance. There are five options that may be turned on and off using the labeled push buttons at the bottom of the control panel, described in Section 3.4.3.

- WX: Shows weather systems in the area. This information is updated via data link.
- STA: Shows navigation stations.
- WPT: Shows all waypoints in the area.
- ARPT: Shows airports.
- EMRG: Shows the candidate routes to alternate destinations, the routes are filtered by the APA dial (see Section 3.4.3). Routes shown as dotted cyan lines.

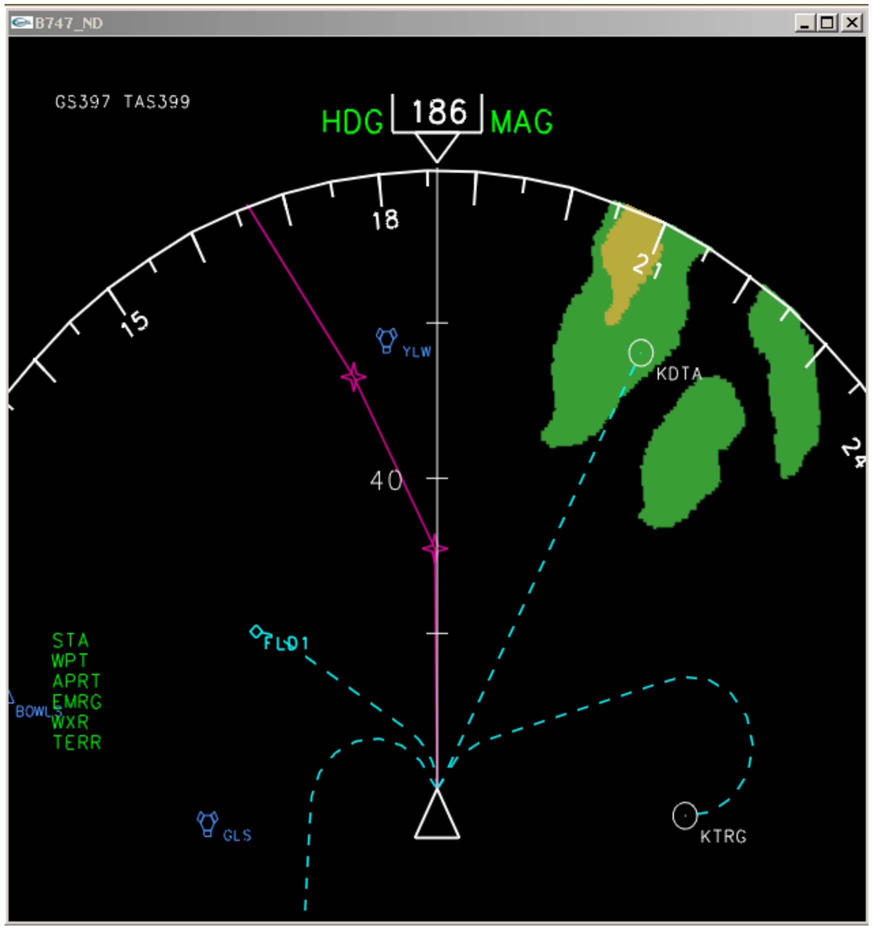


Figure 17: Navigation Display.

3.4.2 Primary Flight Display

The Primary Flight Display (PFD) is located to the left of the ND on the screen in front of the participant, and shown in Figure 18. This display provides information about the current state of the aircraft such as heading, flight speed, altitude, climb/decent rate, and pitch and roll attitudes. Because the participant did not fly by hand, the PFD was provided only as a reference to allow the pilot to be aware of these aspects of the situation. This display provided information which corroborated the emergency that was described, such as loss of altitude as a result of engine failure.

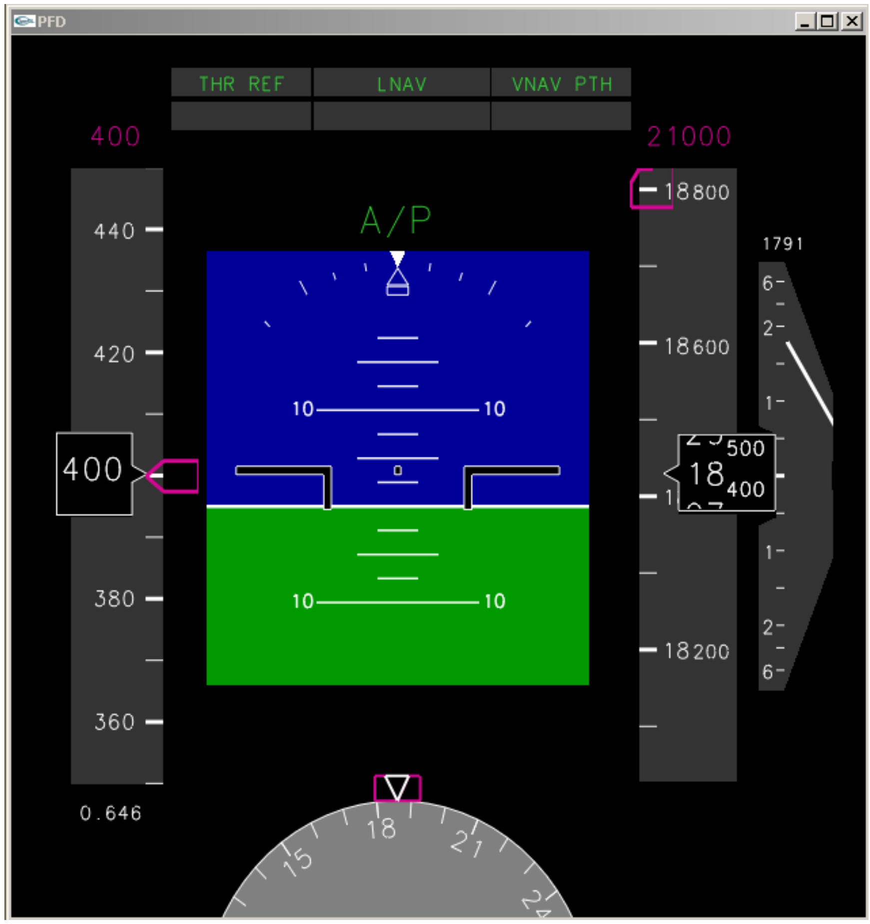


Figure 18: Primary Flight Display.

3.4.3 Navigation Display Control Panel

The Navigation Display Control Panel includes five buttons and two dials, as shown in Figure 19. The buttons allows the five options previously discussed to be toggled on and off. The dial to the right allows the user to set the range (in nautical miles) displayed on the ND. One version of the APA also contained a dial with which the pilot were able to filter possible landing locations. This dial allowed the pilot to quickly indicate the requirements of the landing site.

At each dial position, moving from left to right, more landing sites are shown, but none is removed. The four levels are:

- AIRLINE PRFERD: These are airports where the airline prefers that you divert.

Figure 19: Navigation Display Control Panel.

The airline will be able to perform maintenance, and route passengers on other flights.

- CTRLD ARPTS: These are airports that have control towers. There are no guarantees about the runway length, instrumentation, or emergency services available.
- ALL ARPTS: This filter shows all airports, regardless of runway length, control tower or any other factors.
- ALL FIELDS: This filter shows all potential landing sites, including any flat open space where an aircraft may potentially ditch.

3.4.4 Control Display Unit

The Control Display Unit (CDU) provides a limited subset of the normal CDU functionality. The Route (RTE) and Legs (LEGS) pages, shown in Figures 20 and 21, provide information about the currently planned FMS route. The Alternates (ALTN) page has been redesigned in order to provide more information and support more effective use. The destination options, as filtered by the dial, are ranked according to the overall scores for each potential landing site, as shown in Figure 22. These are the same destination options which are displayed graphically on the navigation display (ND) and may include more than four destinations, in which case the NEXT button is used to move further down the list. This page allows the pilot to see further information about each of the options. This information includes:

- Time to reach site, in minutes
- Distance to site, in nautical miles
- Fuel remaining upon arrival, in pounds
- Length of runway, in feet
- Weather at site, qualitatively assessed as clear, windy, low visibility or thunderstorms
- Medical services available, qualitatively assessed as excellent, good, some or none
- Maintenance available, qualitatively assessed as excellent, good, some or none

Integrating these landing sites into the exiting Alternate Destinations page allowed the pilot to select among options in the same manner which is currently available. After selecting one of the destinations on the list, the pilot was able to view more information about it by pressing MORE INFO (Line 6L). This provided information about the airport and the values that are used by the ranking system for each of the criteria, shown in Figure 23. There are two pages of information, the PREV and NEXT buttons are used to switch between them. After selecting a destination, the pilot may press DIVERT NOW (Line 6R) to choose this option as the divert plan.

Figure 20: Control Display Unit, RTE Page.

(The pilot is not required to view more information about an option in order to select it as the divert, the DIVERT NOW option also appears on the main Alternates page.) This brings up the divert page, Figure 24, which provides a summary of the selected plan. The EXEC button at the upper right of the keypad, is also illuminated. Pressing either the EXEC key or EXECUTE (Line 6R) on the divert page sets the selected plan as the active plan in the FMS and provide autopilot and/or flight director commands for the execution of this plan.

3.4.5 Alert Display

The final display is an alert indicator. It contains a text area which provides a message identifying the alert and a flashing red light to draw attention to the alert. This display is shown in Figure 25.

Figure 21: Control Display Unit, LEGS Page.

3.5 Lab Setup

An external PVC frame was constructed to contain all experiment hardware. The overall setup is shown in Figure 26. This external frame is covered with black cloth in order to block other light sources and isolate the experiment setup. Inside the PVC, a metal and wood frame was constructed. This frame contains two pieces. First, the large shelf in front of the pilot supports the primary flight monitor and computer. The second piece serves as the center console separating the captain's and first officer's seats. The touchscreen was placed in this center console in order to allow the pilots to interact primarily with the CDU near its normal position.

The subjects are seated to the right of the center console in the first officer's seat. Posters were included to provide the pilot with the look and feel of an actual cockpit. These images included the captain's seat, other displays which were not simulated, and a view through the windscreen.

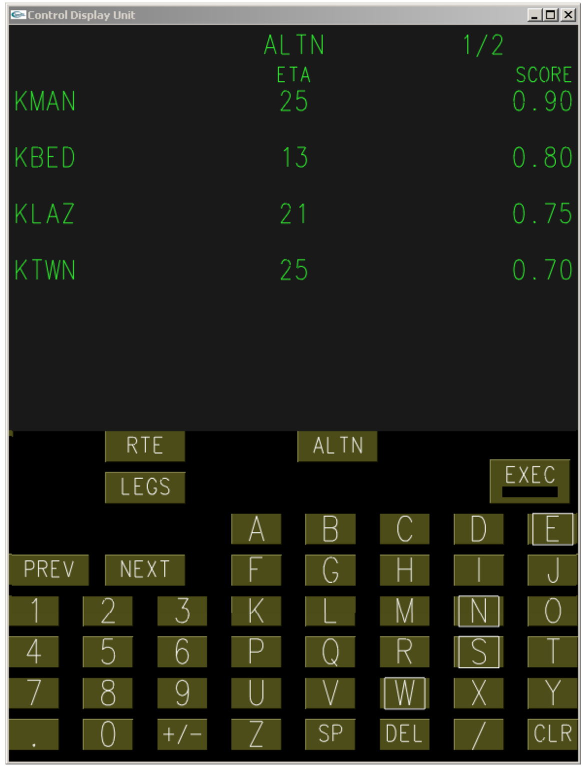


Figure 22: Control Display Unit, ALTN Page.

The computer screen in front of the pilot is an LCD screen (not touchscreen). This screen shows the Primary Flight Display on the left and Navigation Display on the right. The screen on the center console is an LCD touchscreen, where the pilots are able to interact with the system. This screen contains the Control Display Unit, the ND control panel, an alert display and the simulator controls, as shown in Figure 27.

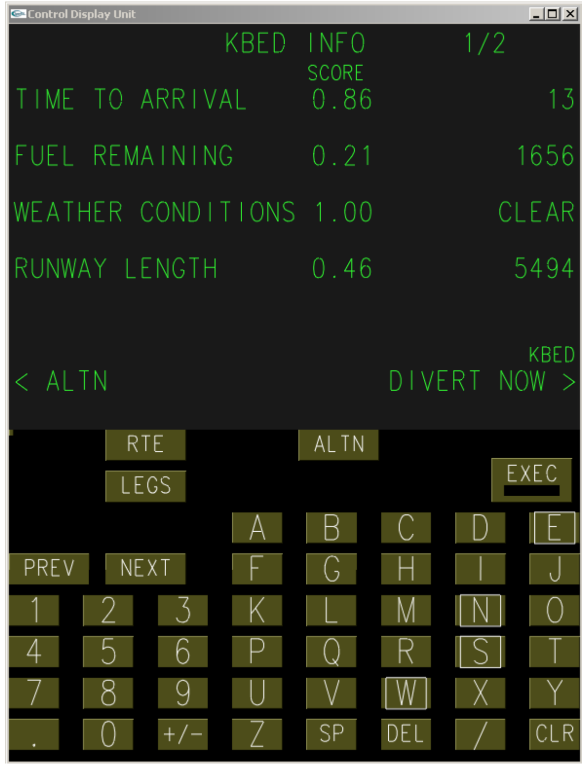


Figure 23: Control Display Unit, MORE INFO Page.

Figure 24: Control Display Unit, DIVERT Page.

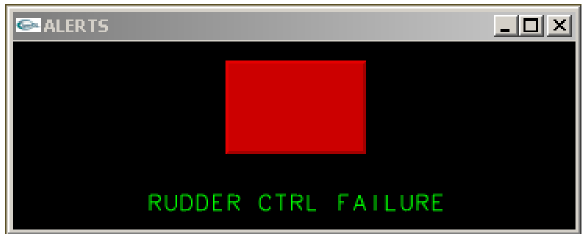


Figure 25: Alert Display.

Figure 26: Frame Surrounding Simulator.

Figure 27: Center Console Screen.

CHAPTER IV

EXPERIMENT DESCRIPTION

This experiment tested for differences in performance for pilots using two variations of an Automated Planning Aid. One variation included the aforementioned dial which facilitates the filtering of landing sites, the other did not include such a dial. The two variations were otherwise identical. In each run, the participant was presented with a scenario in which an emergency occurred. Emergencies that the pilot was expected to have been trained to handle as well as unfamiliar emergencies were presented. The pilot had the opportunity to use the aid, either with or without the dial, to consider the possible alternate landing sites and finally select a plan to land. The pilots did not actually fly the simulated aircraft, and the simulation run ended when the subject had executed a route to the selected landing site in the CDU. Participants were asked to complete a questionnaire after the completion of each run as well as a final set of questions after completing the experiment.

4.1 Scope

In order to ensure that the tests conducted could be completed in a reasonable amount of time, there are a few factors that were not directly addressed. First, pilots did not communicate with anyone during the experiment. In a real emergency situation, the pilots would communicate with other crew members, air traffic controllers and airline dispatchers. While isolating the pilot may not be realistic in some scenarios, pilots may not always receive substantial help from other personnel in all scenarios. For instance, once an emergency has been declared, air traffic controllers may be occupied clearing other traffic from the emergency aircraft's course, and unable to assist in the development of a plan. More simply, an electronics failure may render the aircraft unable to communicate. Additionally, in a highly stressful environment, pilots may focus intently on the task and need to work without distractions, even if these distractions provide useful information. Therefore, any potential aid would need to be self-contained and not rely on the ability to communicate.

In practice, pilots often fly familiar routes, between familiar airports. In the event of an emergency, this familiarity may be useful to quickly recall appropriate landing sites in the area. Unfortunately, the impact of this knowledge of the area was not included in the experiment. In order to fairly compare results, it was necessary that all pilots have similar knowledge of the area. The only means by which this could be assured is by creating an area with which no pilot has familiarity. Therefore, all identifiers used in this experiment were meant to be fictional and any similarity to real places was coincidental. This provided a level playing field for making comparisons across pilots, however, it did discount an important factor.

In an actual emergency situation, the pilot may receive a wide variety of inputs about the situation. These may be as simple as a report from a crew member about a medical emergency. However, it may be a number of visual, auditory or tactile clues. The pilot uses these clues to create a mental picture of the situation. The re-creation of the clues to generate a genuine situation assessment is beyond the scope of this research. Instead, each emergency scenario was described by the captain (not present) who was the pilot flying. This was meant to reproduce for the pilot the type of information which would be gained in an actual emergency. While this did limit the validity of the pilot's situation awareness, it did have the advantage that the pilot's knowledge of the situation could be held constant across different pilots running the same scenario.

4.2 Experiment Design

The experiment included the variation of three independent variables (IV). (Each independent variable had two levels, creating the eight configurations shown in Table 4.2.) These variables were the aid type, familiarity of the emergency, and the quality of the ranking system; each is described in greater detail below. The eight scenarios are listed in Table 3 and are described in Section 4.3. Each participant completed runs using each of the eight configurations paired with the eight scenarios, using the full factorial design shown in Table 4. The design was blocked on Aid Variation, that is, each participant saw four emergencies with one variation, then four with the other variation. A no failure scenario was included in order to attempt to reduce the pilots' expectancy of an emergency. Configuration N is used for the no failure scenario and used the same Aid Variation as the previous run, which varied between participants.

Table 2: List of Configurations.

Configuration	Aid Variation	Familiarity	Ranking System
A	no Dial	Familiar	Optimal
B	no Dial	Familiar	Non-optimal
C	no Dial	Unfamiliar	Optimal
D	no Dial	Unfamiliar	Non-optimal
E	with Dial	Familiar	Optimal
F	with Dial	Familiar	Non-optimal
G	with Dial	Unfamiliar	Optimal
H	with Dial	Unfamiliar	Non-optimal
N	NA	NA	NA

4.2.1 Independent Variables

4.2.1.1 Aid Type

The scenarios were run with two variations of the APA. In one mode, the pilot was able to use the previously described dial to filter possible landing sites. In the second mode, such a dial was not available. The aids were otherwise identical. The primary

Table 3: Scenario Numbers.

Scenario Number	Emergency	Phase of Flight
1	Engine Failure	Climb
2	Engine Failure	Cruise
3	Low Fuel	Cruise
4	Cargo Fire	Descent
5	Rudder Failure	Climb
6	Rudder Failure	Descent
7	Elevator Failure	Cruise
8	Aileron Failure	Cruise

Table 4: Full Factorial Design of Experiment.

Subject	Run 1	Run 2	Run 3	Run 4	Run 5	Run 6	Run 7	Run 8	Run 9
1	A1	C5	N	D7	B3	F4	E2	G6	H8
2	C7	D6	B2	N	A3	E4	G8	H5	F1
3	D8	B4	A2	C6	N	G5	H7	F3	E1
4	B1	A4	C8	D5	H6	N	F2	E3	G7
5	F4	E2	N	G6	H8	A1	C5	D7	B3
6	E4	G8	H5	N	F1	C7	D6	B2	A3
7	G5	H7	F3	E1	N	D8	B4	A2	C6
8	H6	F2	E3	G7	B1	N	A4	C8	D5

focus of this research was to determine the value of the addition of such a dial.

4.2.1.2 Familiarity

The evaluation scenarios simulated two general types of emergencies: familiar emergencies and unfamiliar emergencies. In the familiar scenarios, the aircraft's performance was either unaltered or was altered in a manner that pilots have been trained to handle, such as a single engine failure. The second type of scenario was a performance altering emergency in which the failure was one which the pilot has not been trained to manage, such as a stuck elevator. The scenarios were designed such that each was comparable in difficulty and number of options that must be considered. The comparison of these two categories of emergency is important because pilots may make decisions differently in a familiar scenario than they do in an unfamiliar one. See Section 4.3 for specific information about the emergency scenarios which were used.

4.2.1.3 Automation Accuracy

As previously mentioned, the prioritization scheme was not directly studied in this research. One of the difficulties of developing such a scheme is the highly variable nature of airborne emergencies. Therefore a proposed prioritization scheme will likely be brittle and only valid within a certain scope of emergencies. This is an important factor to consider because the pilot's use of the aid will likely differ depending on how appropriate the prioritization output is to the scenario. For this experiment, it was assumed that the "best" landing site was be determined by an expert after taking into account all available information. A realistic, codable scheme will likely rank this "best" site highest in emergencies for which it was designed and rank the "best" landing site lower in other emergencies. For this reason, the prioritization output that was used also was imperfect in some runs. This variable had two levels, an optimal ranking system and a non-optimal ranking system. After the sites had been scored by an expert (see Section 3.2.1.2, this ranking was used when the ranking system used was optimal. The scores and rankings were changed from this optimal ranking in order to produce a non-optimal ranking, which guaranteed that the highest ranked site in the optimal scheme was ranked lower in the non-optimal scheme.

4.2.2 Dependent Variables/Measurements

Two primary metrics were considered. First, the pilot's ability to choose the "best" landing site. An expert was consulted in order to provide aggregate scores for each landing site based on all information available in the scenario. These scores served to rank each of the landing sites in each scenario. The second metric was the amount of time that pilots spent in the planning process. A reduction in time promotes safe flight by allowing the pilots to focus on other important tasks associated with handling the specific emergency, such as crew coordination, and alerting personnel on the ground. The time required for the pilot to select a path was used to determine the efficiency with which the pilot was able to develop a plan. This time was measured from the moment the emergency occurs to the time that the pilot has selected an option in the CDU.

In support of these measures, other secondary measures were used to assess the APA. The number of candidates that the pilot reviewed and the number of times the pilot turned the filter dial (when available) were also measured. In addition to simply comparing the total time to complete the task, these measures allowed for a more granular analysis of the task.

The dependent variables (DV) of interest are listed in Table 4.2.2. Also listed are the levels of each of the variables, which were used for the analysis.

Table 5: Table of Variables.

Measure	Level
Time	continuous
Quality	ordinal
Number of Alternates viewed	discrete
Number of Dial Turns	discrete
Situation Awareness	discrete
Workload	continuous
APA Assessment	discrete

4.2.2.1 Time

This time was measured from the time the emergency occurs to the time that the pilot has executed a path in the CDU.

4.2.2.2 Quality

Each scenario involved a number of potential landing site options. These options were ranked according to the appropriateness for the given scenario. This aspect of the participant's performance was based on the rank assigned to the selected landing site.

4.2.2.3 Number of Alternates viewed

The number of alternates viewed was the number of landing sites for which the pilot viewed the "MORE INFO" page. This was automatically recorded by the system and was the total number of alternates viewed before and after the emergency occurs. The number of alternates viewed was recorded in order to evaluate factors that influenced the amount of time required to make a selection.

4.2.2.4 Number Dial Turns

The number of times that the filter dial was turned was also automatically recorded by the system. This measure may be useful to determine the impact the presence of the dial has on pilots as they complete the task. A large number of dial turns may indicate that the pilot is unable to find a suitable filter which reduces the number of available options without eliminating acceptable landing sites. This measure may also provide some indication of how utility of the filter dial, as a low number may indicate that pilots are not using the dial effectively.

4.2.2.5 Situation Awareness

The participant's situation awareness was assessed immediately after the completion of each run. The displays were blanked and the participant was asked ten questions about the current scenario. The questions assessed all three levels of SA [11, 10]. Level 1 assessed the pilot's perception of cues, level 2 assessed the pilot's comprehension of the situation, and level 3 assessed the pilot's ability to forecast future events. The ten questions contained five level 1 questions, three level 2 questions and two level 3 questions. These were drawn from a pool of twelve level 1, ten level 2 and five level 3 questions. For a complete list of SAGAT questions see Appendix A.

4.2.2.6 Workload

The participants were asked to evaluate the workload experienced in each scenario in order to assess the feasibility of its use in a real emergency. The NASA Task Load Index (TLX) is a tool which is used to assess and record a user's workload [15]. The TLX assessed workload for six different sources of workload; mental demand, physical demand, temporal demand, performance, effort, and frustration.

4.2.2.7 APA Assessment

Upon completion of the experiment, the participants were asked to complete a questionnaire, which included questions about the pilot's experience in each simulated flight, as well as an evaluation of both variations of the APA. These included subjective assessments as well as a rating from one to ten based on the Modified Cooper-Harper for displays [8].

4.3 Scenarios

Each pilot participated in a total of nine runs (eight emergency scenarios and one no failure scenario). Each scenario was characterized by the emergency situation that occurs, the phase of flight during which the emergency took place, the alternate landing sites which were available and the time at which the failure occurred (from the time the simulation was started). Table 4.3.4 provides the specific combination used to create the eight emergency scenarios.

4.3.1 Landing Sites

Each scenario had a fixed number of potential landing sites, dependent on the phase of flight in which the emergency occurred. Emergencies occurring in climb had three sites, cruise had six sites, and descent had four sites. All identifiers of airports, waypoints, and navigation stations were fictional. The decision to use fictional places is discussed in more detail in Section 4.1.

4.3.2 Emergencies

There were a total of six types of emergency situations, three familiar and three unfamiliar. In each case the captain, who was the pilot flying, described the emergency. The flight display also showed any appropriate changes (such as descent) and the alert display annunciated the appropriate message.

4.3.2.1 Familiar Emergencies

The three familiar emergencies used were engine failure, low fuel, and fire onboard. In the case of the engine failure, the captain said, "Oh, we've got a right engine failure! I've completed the checklists and I'am able to hold the aircraft in steady flight, but I can't maintain altitude. Call ATC and declare an emergency, engine failure, request lower, unable to maintain altitude." The alert display said, "LEFT ENGINE FAILURE." In the case of the low fuel emergency, the captain said, "The fuel gauge is reading extremely low. I can't determine the correct amount of fuel remaining. It may be incorrect now, or it may have been incorrectly loaded before the flight. The appropriate checklists have been completed, but we cannot be sure how much fuel we have remaining. Contact ATC to make them aware of the situation." The alert display said, "LOW FUEL." In the case of the fire emergency, the voice of another crew member first said, "We had a fire in the cargo area!! The fire seems to be extinguished, but there is still a large amount of smoke." Then the captain said, "OK, we've got a fire in the cargo area. It seems to be extinguished but there is still a large amount of smoke. The appropriate checklists have been completed. Call, ATC, declare an emergency. Inflight fire, request lower and nearest airport." The alert display said, "FIRE: CARGO HOLD."

4.3.2.2 Unfamiliar Emergencies

In order to allow for appropriate descriptions and understanding of failures, flight control failures for which pilots have not been trained were used as the unfamiliar emergencies. These were stuck rudder, stuck aileron, and stuck elevator. In the case of the stuck rudder, the captain said, "The rudder seems to be stuck. I have no yaw control. I've completed the appropriate checklists, but we need to declare an emergency." The alert display said, "RUDDER CTRL FAILURE." In the case of the stuck aileron, the captain said, "The right aileron is stuck! I am able to compensate and keep the aircraft level, but I'm having a very hard time turning. The appropriate checklists have been completed. We need to declare an emergency." The alert display said, "AILERON CTRL FAILURE." In the case of the stuck elevator, the captain said, "The elevator seems to be stuck. I cannot keep the nose up without full throttle. The appropriate checklists have been completed. We need to declare an emergency." The alert display said, "ELEVATOR CTRL FAILURE."

4.3.3 Phase of Flight

The eight emergencies each occurred in one of three phases of flight; climb, descent or cruise. The emergencies which were repeated occurred in different phases of flight each time.

4.3.4 Time to Emergency

The amount of time which passed after the simulation began until the emergency occurred was different for each scenario. The amount of time was not less than 45 seconds and not more than 105 seconds.

Table 6: Table of Scenarios.

Familiarity	Emergency	Phase of Flight	Time to Emergency (sec)
Familiar	Engine Failure	Climb	85
Familiar	Engine Failure	Cruise	105
Familiar	Low Fuel	Cruise	60
Familiar	Cargo Fire	Descent	45
Unfamiliar	Rudder Failure	Climb	60
Unfamiliar	Rudder Failure	Descent	45
Unfamiliar	Elevator Failure	Cruise	90
Unfamiliar	Aileron Failure	Cruise	105

4.4 Procedure

4.4.1 Briefing and Training

Before entering the simulator, each participant read a briefing document. This document introduced him to the features of the simulator and the procedures which would be used to conduct the experiment. The entire briefing document can found in Appendix B.

After reading the introductory material the participant entered the simulator to complete at least two practice scenarios. In the first training scenario, no failure occurred. This run simply allowed the pilot to explore the interface and familiarize himself with the tools available. The second run presented a simple engine failure scenario. This allowed the pilot to gain an expectation of how emergencies would be presented and how the tools would allow him to plan a diversion. After each run the participant was given sample SAGAT questions to familiarize him with the format and types of questions which would be asked. This was done primarily in order to avoid any differences between the first couple of runs and the later runs in terms of answering SAGAT questions. The participant was also asked to complete the TLX workload questions. The participant was then given the option to run either of the training scenarios again or to begin with the test runs.

4.4.2 Test Runs

After completing the training runs, each pilot completed nine test runs. Eight of these runs contained one of the aforementioned emergencies and one was completed without an emergency. The no emergency run was included in order to slightly reduce the pilots expectancy of an emergency. The pilot joined each flight in progress and was able to use the tools available to gain an understanding of the situation. Between 45 and 105 seconds after the simulator was started some emergency situation was presented by the captain, through a recorded message. The participant then used the tools to determine the best landing site for the given situation.

After executing a route to the selected landing site, the simulator was closed and the participant was handed a clipboard and asked to complete the ten SAGAT questions. After completing the SAGAT questions, the participant was asked to complete the TLX questionnaire on the touchscreen.

4.4.3 Post Test Questionnaire

Upon completion of all nine runs, the pilot was given an additional set of questions pertaining to the experiment as a whole. These included subjective questions about the features and usage of the APA and the Modified Cooper-Harper ranking sheet. These questions can be found in Appendix C.

After the completion of the questionnaire, the pilot was asked to evaluate a modified version of the CDU in which the aforementioned Update Weights page was made available. The participant was asked to comment on the usefulness of the ability to manipulate the criteria weights in general, as well as the specific interface presented.

4.4.4 Deception

The recruitment and introductory material informed participants that the compensation they would receive was dependent on how well they performed the emergency planning task (but was not dependent on answering the SAGAT and TLX questions).

After the completion of all runs, the participants were made aware of the deception and that all participants would receive the entire sum of compensation. The Disclosure of Deception document is provided in Appendix D.

CHAPTER V

ANALYSIS OF RESULTS

5.1 Demographics

A total of eight pilots participated the experiment procedure. These pilots all hold an Airline Transport Pilot certificate and are experienced in a variety of aircraft, primarily aircraft manufactured by Boeing, Airbus and Bombardier. One participant is recently retired, while the others currently fly with an airline. Figure 28 shows the amount of experience that the participants have in flight hours. The average number of flight hours was 8194 hours.

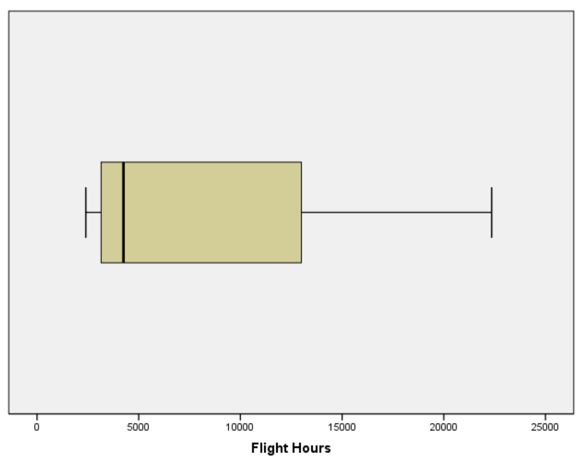


Figure 28: Distribution of Flight Experience

5.2 Performance

The pilots' performance was measured in terms of two parameters. First, the pilots' time to complete the planning task was measured to ensure that pilots act quickly. In the event of an emergency, delays in selecting a divert plan can be costly. Also, within this experiment, the replanning task was the subject's only responsibility, unlike an actual onboard emergency, where the pilot would have his attention divided between a number of important tasks. Therefore, a reduction in time spent selecting a divert plan allows more time for the pilot to attend to other tasks.

In addition to measuring the time required to complete the task, the second measure of performance was the quality of the landing site which was selected. While making a decision quickly is very important, this speed must not come at the expense of making a good decision. As previously discussed in Section 4.2.2.2, the possible landing sites were ranked in order of how appropriate each site was for the particular emergency situation encountered in a run. The ranking of the selected site was used to measure the quality of the landing site. This ranking is appropriate because it does not take into account how appropriate a site is, only how appropriate it is among the options available in a given scenario.

Overall, the average time required to select a landing site was 109.8 seconds, with a standard deviation of 47.8 seconds. The uniqueness of each scenario is an important factor which largely accounts for the difficulty in selecting a landing site. Each test run was completed in a suitable amount of time with the exception of one test run by one subject in which the engine failed during climb. The subject in this run selected an alternate landing site 217.9 seconds after the failure occurred, which was considered too long by the expert, given the aircraft's altitude at the time. All other test runs by this subject and all other subjects were completed in an acceptable amount of time. As shown in Figure 31(a), the average times vary between scenarios. The median times were also varied. The shortest median time was 58.0 seconds for the fire emergency and the longest median time was 186.7 seconds for the aileron control failure emergency. While there was some variation between all scenarios, the times for the aileron failure scenario were significantly higher than for other scenarios. (The average time of the aileron failure scenario was 206.8 seconds compared to an average of 96.0 for all other scenarios.)

The ability of pilots to select the best sites was generally quite good. The average ranking of landing sites selected was 1.78. (Because this measure is ordinal, the median would be a more appropriate measure, however because the highest ranked landing site was selected in over half of the runs, the median is 1 and does not provide much information.) The selected landing site was ranked highest in 57.8% of runs. Figure 29 shows the distribution of the ranking of landing site selected in all runs.

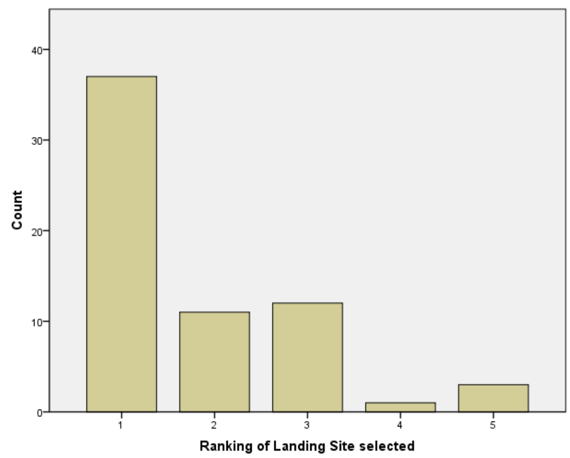


Figure 29: Distribution of Landing Site Quality.

This measure also varied by each emergency scenario. Figure 31(b) demonstrates the results of the landing site quality measure for each scenario. These results show that the difference in overall appropriateness may have not have been the same for each scenario, as was intended. For the engine failure during climb and the rudder control failure in climb, every pilot was able to select the most appropriate landing site. For fire emergency scenario, only one participant selected the most appropriate landing site. This may not be surprising, considering that time is most limited in the event of an onboard fire. For the elevator failure scenario, a number of different sites were selected, which may indicate that the differentiation between the best landing site and the other landing sites may not have been substantial enough.

In each scenario, there may be some trade-off between the two measures of performance, time and quality. It was expected that some pilots may spend more time deliberating and coming to a better decision, while others may make decisions more quickly at the expense of the quality of the decision. The Spearman's signed rank test was used to determine the correlation between these two measures. The results showed that there was a positive correlation between time and quality, r_s = .203, p = .107. Figure 30 shows that runs in which the pilot spent more time before making a decision resulted in slightly poorer decision quality.

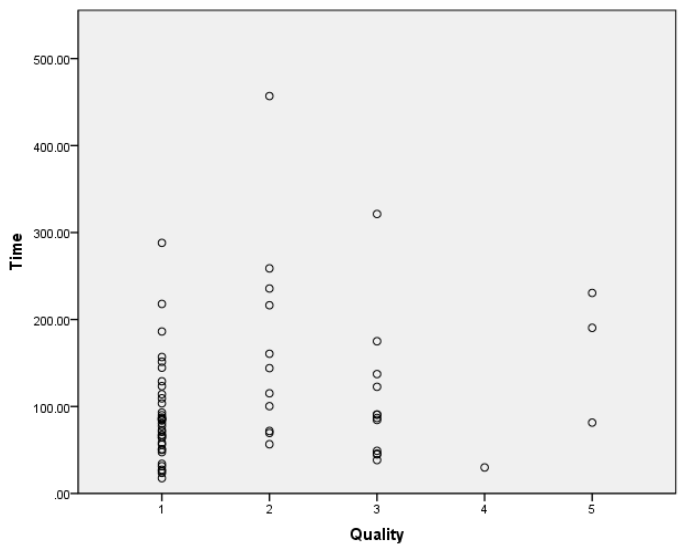
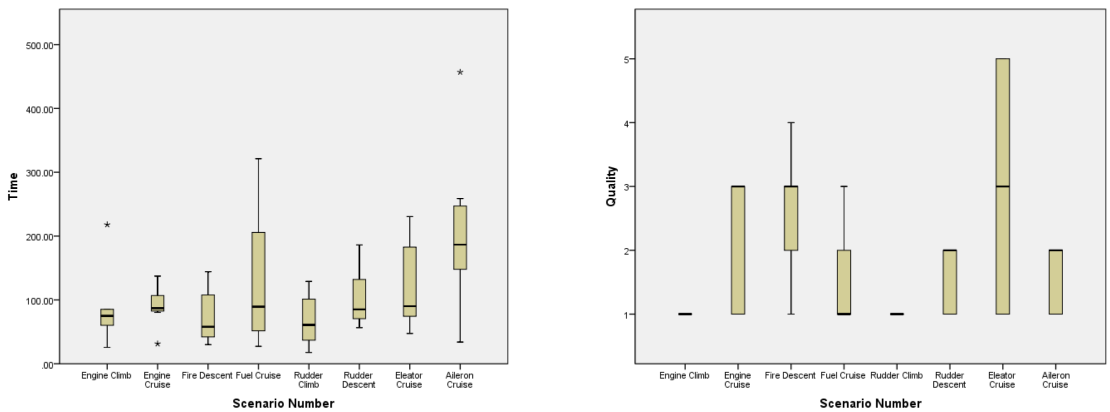


Figure 30: Correlation between Time and Decision Quality.

All of the statistical tests which were conducted used 95% confidence intervals (α = .05). However, this criteria reports many of the results as not statistically significant. The non-significance of these results is not unexpected given that the observed powers generally lower than the common 0.80 standard for adequacy. These powers are a result of small difference in means compared to the variations. Coupled with small sample sizes, this results in difficulty determining whether a significant effect was observed. For this reason, an emphasis on trends and observations is made despite results which are generally not statistically significant.

To better account for differences between scenarios, the analysis was run in two ways. First, the raw results were used to test for the effects of the independent



(a) Time to Complete by Scenario.

(b) Landing Site Quality by Scenario.

Figure 31: Performance Results.

variables. Second, the results were recalculated such that each measure represented the deviation from the mean for the scenario used in the run. For example, the time for a given run was represented as an increase or decrease from the mean time for the scenario used in that run. The analysis was re-run with these modified measures. Though slight differences between the two analyses were observed, the results did not vary significantly. Therefore, the results of the analysis using the un-modified data are presented.

5.2.1 Impact of APA Dial

The primary metric of interest was the effect that the addition of the filter dial had on both aspects of performance, the time to select a landing site and the quality of the landing site selected. The mean time for cases with the filter dial was actually slightly higher than those without the dial (μ_{dial} = 115.3 seconds, μ_{noDial} = 104.4 seconds). Figure 33(a) shows a comparison of the time to complete the task using each of the two variations.

These results must be interpreted in the presence of the other two independent variables, the familiarity of the emergency and the quality of the ranking scheme.

In order to assess the impact of the variables on the time measurement, a three-way Repeated Measures ANOVA (RM-ANOVA) was used. The main effect of APA variation was non-significant, $F(1{,}7) = 0.217$, $p = 0.655$. The main effect of the ranking scheme used was also non-significant, $F(1{,}7) = 0.007$, $p = 0.938$.

The effect of the familiarity of the emergency on time to complete was reported as significant, $F(1{,}7) = 6.979$, $p = 0.033$. As anticipated, the time to select an alternate was lower in the case of familiar emergencies than in cases of unfamiliar emergencies ($\mu_{fam} = 94.9$ seconds, $\mu_{unfam} = 124.7$ seconds). The impact of the dial was different in the familiar cases than in the unfamiliar cases. In familiar emergencies, the addition of the dial resulted in a large increase in the average time from $\mu_{noDial} = 79.3$ seconds to $\mu_{dial} = 110.6$ seconds. However, in unfamiliar emergencies, the dial reduced the average time from $\mu_{noDial} = 129.5$ seconds to $\mu_{dial} = 119.9$ seconds. This comparison is shown in Figure 32. The interaction of these effects was also reported non-significant, $F(1{,}7) = 1.893$, $p = .211$.

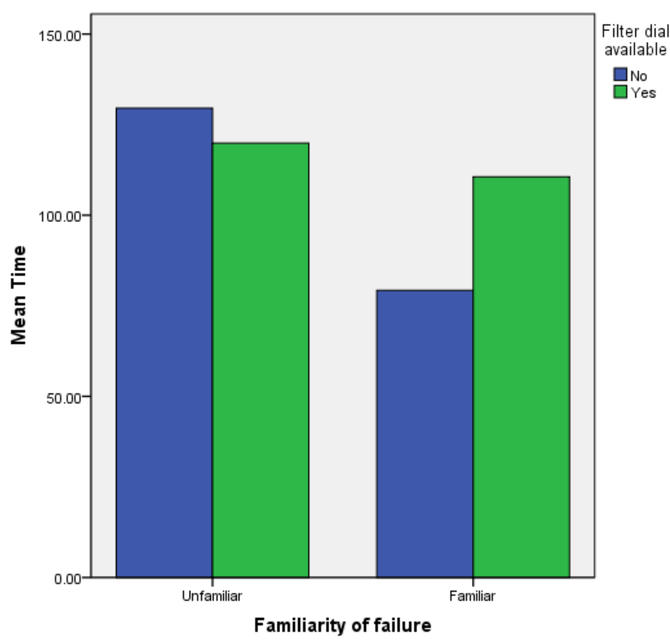
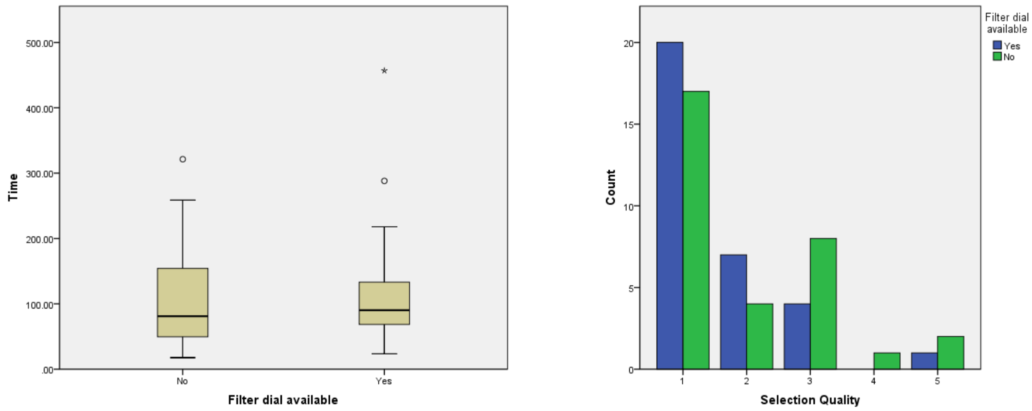


Figure 32: Impact of APA Variation and Familiarity on Time.

As shown in Figure 31(a), the aileron failure scenario has a mean time to complete



(a) Time to Complete

(b) Quality of Selection

Figure 33: Effect of Dial on Performance

which was much higher than any other case. In order to investigate this disparity, a one-way RM-ANOVA was used with scenario as the independent variable and time as the dependent. Mauchly's Test of Sphericity showed that the assumption of sphericity was violated, $\chi^2(27) = 51.555$, $p = .011$. The degrees of freedom are corrected using the Greenhouse-Geisser estimate of sphericity ($\epsilon_{GG} = .285$). The effect of scenario was reported as significant, $F(1.997, 13.978) = 4.745$, $p = .027$. Simple contrasts showed that the time to complete for each of the other seven scenarios differed significantly from the aileron failure scenario (all $p < .05$). In order to exclude this scenario from the analysis a Linear Mixed Model was used. This analysis showed that none of the three independent variables had a significant effect on time to complete the task. (For main effect of APA variation, $F(1{,}37.261) = .579$, $p = .452$; for main effect of familiarity, $F(1{,}41.873) = .023$, $p = .879$; for main effect of ranking system, $F(1{,}42.713) = 1.509$, $p = .226$.) Upon, exclusion of this scenario, the impact of the dial in unfamiliar scenarios is not dramatically altered. The dial reduced the average time from $\mu_{noDial} = 105.830$ seconds to $\mu_{dial} = 97.346$ seconds.

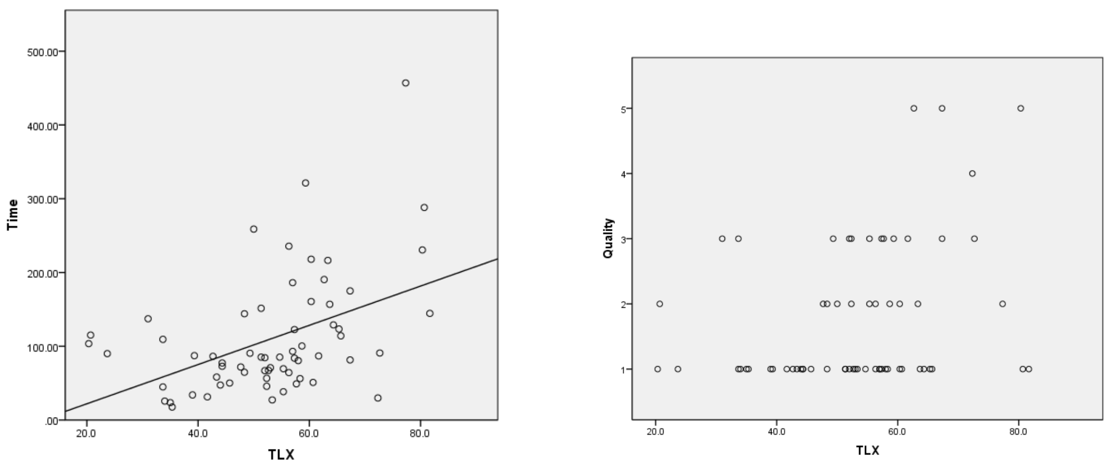
The impact of the APA variation was also tested with regards to the quality of the landing site selected. Cases in which the filter dial was present resulted in slightly better quality of alternates selected. Figure 33(b) shows the results of rank of landing site selected based on the presence of the dial. The selected landing site was ranked first or second in 84.4% of cases which included the dial, compared to only 65.6% of cases which did not include the dial. Similarly, the site was ranked third or lower was 34.4% for cases without the dial, compared to only 16.6% for cases with the dial present. Because this dependent variable is ordinal, the Friedman's ANOVA test was used. Unlike the RM-ANOVA, the Friedman test does not allow for the testing of interactions between independent variables. The results of the Friedman test showed that the difference in quality of landing site selection between variations of the aid was not significant, $\chi^2(1) = .727$, $p = .394$. The results for both familiarity of emergency and optimality of ranking system used was also reported as non-significant, $\chi^2(1) = .222$, $p = .637$ and $\chi^2(1) = .059$, $p = .808$, respectively.

5.3 Secondary Measures

5.3.1 Workload

As previously discussed, the cockpit environment can be very stressful in the event of an emergency. The pilots have a number of tasks which require their attention and which must be completed in a timely fashion. Therefore the workload that the pilot encounters as a result of the APA system should be minimized. The workload was only slightly reduced by the addition of the dial ($\mu_{dial} = 51.1$, $\mu_{noDial} = 54.9$). A three-way RM-ANOVA test was used to determine the impact that the three independent variables had on the pilots' workload. The impact of the addition of the dial was reported as non-significant, $F(1, 7) = 1.878$, $p = .213$. The levels of familiarity and ranking system resulted in little change in the workload. The main effects of the familiarity and ranking system were therefore not significant, $F(1, 7) = 1.277$, $p = .296$ and $F(1, 7) = 0.622$, $p = .456$ respectively.

The addition of the dial was not shown to reduce the pilots' workload. However, a reduction in workload, regardless of the source, may result in better performance.



(a) Time to Complete (b) Quality of Selection

Figure 34: Scatter plots of Workload Score versus Performance.

Therefore a correlation analysis was used to determine if there was indeed a correlation between pilots' workload and their performance. The Spearman's rank correlation coefficient was used to determine that the correlation between the workload and time $r_s = .432$, $p < .001$ as well as between workload and quality $r_s = .286$, $p = .022$ were both significant. Figure 34 shows the relationships between workload and performance.

5.3.2 Situation Awareness

In order to make a good decision, the pilot must be aware of the situation at hand. The results of the situation awareness questionnaire were used to determine the impact of the aid variation on the pilots' understanding of each situation. Overall, the pilots' performance on all situation awareness questions was lower than expected. Figure 35 shows the percentage of correct questions answered at each of the three levels. These low percentages of correct responses may be partly attributed to the questions which were asked. A number of the questions required that pilots recall airport identifiers in order to correctly answer the questions. A few participants complained that the use of unfamiliar and meaningless identifiers made them very hard to remember. One pilot described the unfamiliar identifiers as "alphabet soup." Also, pilots were able to keep only the important identifiers in mind, and these were kept only long enough to complete the scenario. Upon completion of each run, pilots "data dumped" and no longer recalled meaningless identifiers. These factors likely contributed to low numbers of correct responses.

Despite the low scores, three separate three-way RM-ANOVAs were used to test for effects of the independent variables, one for each level of situation awareness (SA). In general, though not necessarily significant, the addition of the dial increased situation awareness. Situation awareness was also generally better in familiar emergencies compared to unfamiliar, and better in cases which used the optimal ranking system as opposed to those which used the non-optimal system. The only deviation from this trend was Level 2 SA, which was better in unfamiliar emergencies than familiar ($\mu_{fam} = 27.5\%$, $\mu_{unfam} = 31.3\%$). However, this effect was not found to be significant, $F(1{,}7) = 1.615$, $p = .244$.

The addition of the dial increased Level 1 SA from $\mu_{noDial} = 48.1\%$ to $\mu_{dial} = 58.8\%$, which was marginally significant $F(1{,}7) = 3.781$, $p = .093$. Level 1 SA was also better for cases with familiar emergencies than for cases with unfamiliar emergencies ($\mu_{fam} = 58.8\%$, $\mu_{unfam} = 48.1\%$). This effect was also reported as significant $F(1{,}7) = 8.190$, $p = .024$. Level 2 SA was not impacted by any single effect, but was significantly impacted by the interaction of the dial both with familiarity and with ranking system, $F(1{,}7) = 9.000$, $p = .020$ and $F(1{,}7) = 11.200$, $p = .012$ respectively. There was no significant effect on Level 3 SA. The full results are reported in Table 5.3.2.

As with workload, changes in situation awareness, regardless of the source, may have an impact on the pilots' performance. In order to test for this correlation, the Spearman's rank correlation coefficient test was again used. The results show that there was no significant correlation between performance and Level 1 or Level 2 SA. There was however a marginally significant negative correlation between Level 3 SA

Table 7: Impact of APA Variation, Familiarity and Ranking System on Levels 1, 2, and 3 SA

(a)

SA Level 1

Source	df, Err. df	F	p	Observed Power
Dial	1,7	3.781	.093	.390
Familiarity	1,7	8.190	.024	.691
Ranking System	1,7	2.166	.185	.247
Dial × Familiarity	1,7	.549	.483	.099
Dial × Ranking System	1,7	.006	.938	.051
Familiarity × Ranking System	1,7	.008	.930	.051
Dial × Familiarity × Ranking System	1,7	1.163	.317	.155

(b)

SA Level 2

Source	df, Err. df	F	p	Observed Power
Dial	1,7	.636	.451	.107
Familiarity	1,7	1.615	.244	.197
Ranking System	1,7	1.378	.279	.175
Dial × Familiarity	1,7	9.000	.020	.731
Dial × Ranking System	1,7	11.200	.012	.818
Familiarity × Ranking System	1,7	1.000	.351	.140
Dial × Familiarity × Ranking System	1,7	.583	.487	.098

(c)

SA Level 3

Source	df, Err. df	F	p	Observed Power
Dial	1,7	.283	.612	.075
Familiarity	1,7	.035	.857	.053
Ranking System	1,7	1.224	.305	.161
Dial × Familiarity	1,7	2.882	.133	.312
Dial × Ranking System	1,7	.164	.697	.064
Familiarity × Ranking System	1,7	.042	.844	.054
Dial × Familiarity × Ranking System	1,7	.247	.634	.074

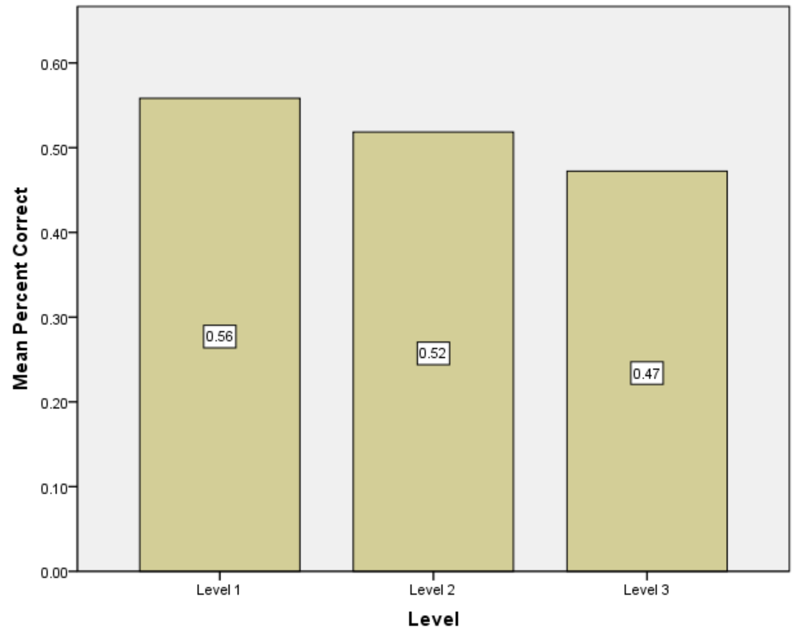


Figure 35: Percentage of SA Questions Answered Correctly

Table 8: Correlations Between Situation Awareness and Performance.

	Time		Quality	
	r_s	p	r_s	p
Level 1 SA	-.102	.425	-.084	.507
Level 2 SA	-.063	.624	-.096	.448
Level 3 SA	-.212	.092	-.196	.120

and the time to select a landing site r_s = -.212, p = .092. The full set of correlations is shown in Table 5.3.2. In addition to the correlation between situation awareness and performance, there was a significant correlation between situation awareness and workload, r_s = -.307, p = .009 for Level 1 SA, r_s = -.212, p = .073 for Level 2 SA, and r_s = -.187, p = .115 for Level 3 SA. This implies that runs in which the subject had lower situation awareness, his workload was also higher. That is, subjects who had a better understanding of the situation encountered a lower workload.

5.3.3 Number of Solutions Viewed

In order to understand possible causes which may lead to better performance in certain combinations of the independent variables. Two such causes were considered; first, the number of landing sites for which the pilot viewed more information, and second, the number of times the dial setting was changed, for cases in which the dial was available. Certainly, viewing more options may become a time-consuming task, however, it may lead to a better understanding of the alternates available and thus a better selection quality. A three-way RM-ANOVA was conducted on the number of landing sites viewed to determine the impact that the independent variables had on this measure. The average number of solutions viewed was slightly lower for cases with the dial than for those without the dial ($\mu_{dial} = 4.5$, $\mu_{noDial} = 5.0$). This difference was reported as non-significant $F(1{,}7) = .356$, $p = .569$. The ranking system used had very little impact on the number of solutions viewed, resulting in a non-significant difference, $F(1{,}7) = .028$, $p = .871$. However, the familiarity of the emergency had a significant impact on the number of solutions viewed, $F(1{,}7) = 13.139$, $p = .008$. The average number of alternates viewed in familiar emergencies was lower than cases with unfamiliar emergencies ($\mu_{fam} = 3.6$, $\mu_{unfam} = 5.9$). In familiar emergencies, the addition of the dial slightly increased the number of solutions viewed from $\mu_{noDial} = 3.5$ to $\mu_{dial} = 3.7$. However, in unfamiliar emergencies, the dial reduced the average number of solutions viewed from $\mu_{noDial} = 6.5$ to $\mu_{dial} = 5.3$. This comparison is shown in Figure 36. The interaction of these effects was also reported non-significant, $F(1{,}7) = .655$, $p = .445$.

As with the time to complete measure, the number of solutions viewed differed greatly between scenarios. As Figure 37 shows, the aileron failure scenario again stands out, with an average number of solutions viewed of $\mu_{aileron} = 10.3$, compared to an overall average of $\mu_{all} = 4.5$. To verify this significance, a one-way RM-ANOVA was used to determine if the effect of scenario was significant on the number of solutions

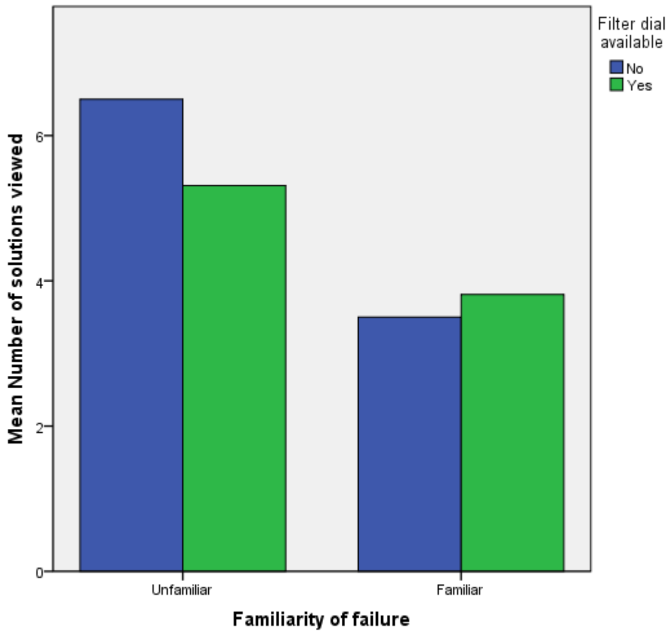


Figure 36: Impact of APA Variation and Familiarity on Number of Solutions Viewed.

viewed. This effect was indeed significant, $F(7{,}49) = 6.484$, $p < .001$. Again a linear mixed model was used to test the effects after excluding runs with the aileron failure scenario. These results showed no statistically significant differences between levels of each effect. (For main effect of APA variation, $F(1{,}35.655) = .057$, $p = .812$; for main effect of familiarity, $F(1{,}47.752) = .779$, $p = .382$; for main effect of ranking system, $F(1{,}48.537) = .281$, $p = .599$.)

In order to determine the impact that the number of alternates viewed had on the pilots' performance. The Spearman's rank correlation coefficient test was used to determine the correlation between the number of alternates viewed and the time to select an alternate and again with the quality of the landing site selected. There was a significant positive correlation between the number of alternates viewed and the time to complete the task, $r_s = .737$, $p < .001$. The positive correlation between the number of alternates viewed and the quality of the selection made was not significant, $r_s = .171$, $p = .178$. This implies that participants who viewed more solutions took

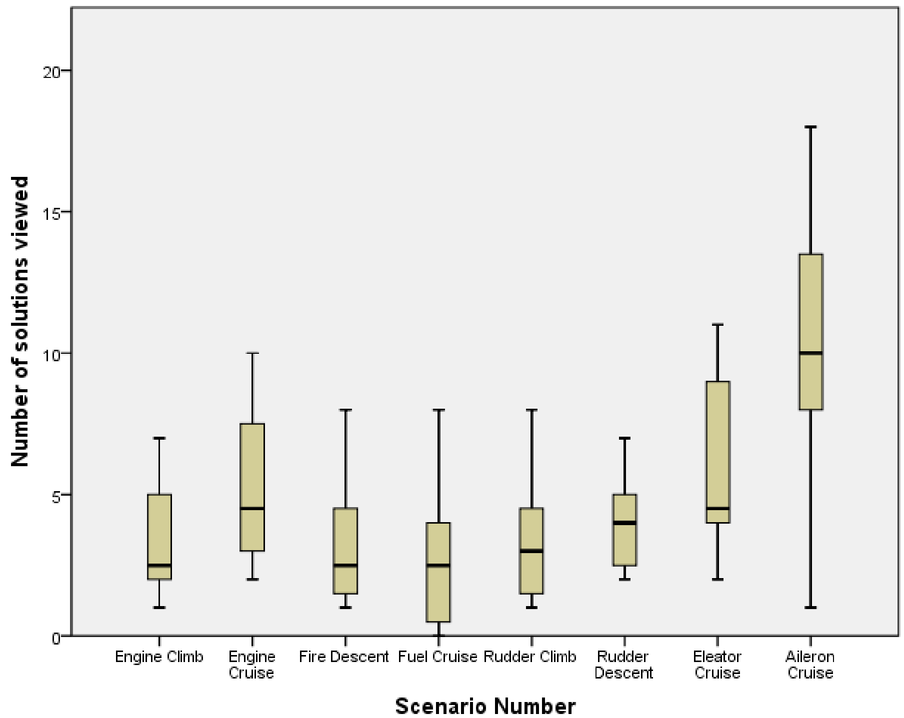


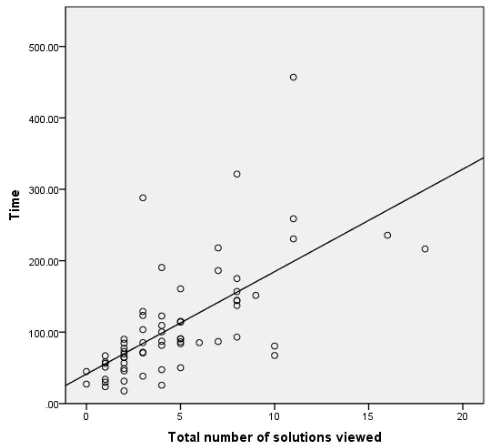
Figure 37: Number of Solutions Viewed by Scenario.

significantly more time to make a selection, but did not make a significantly better decision. These correlations are seen in Figure 38.

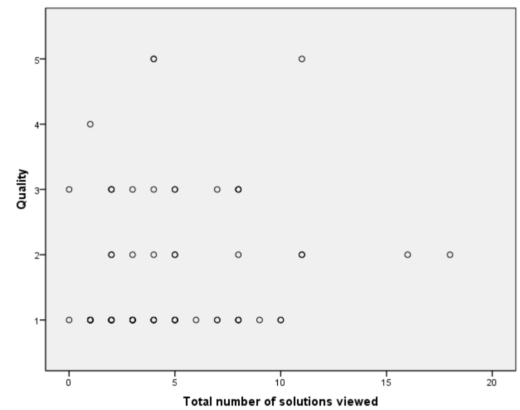
In addition to this correlation, a correlation was also found between the number of alternates viewed and the workload encountered. The Spearman's rank correlation coefficient test was again used to determine that this relationship was reported significant, $r_s = .352$, $p = .002$. This positive correlation implies that runs which pilots viewed more alternates, they also encountered a higher workload. Conversely, when pilots only viewed a few alternates, the workload encountered was lower.

5.3.4 Number of Dial Turns

For those cases in which the filter dial was available, the number of times the dial setting was changed was also a measure of interest. This may indicate some indecision on the part of pilot as to which setting is most appropriate or a lack of satisfaction with the alternates available at a particular setting. A two-way RM-ANOVA test was



(a) Time to Complete

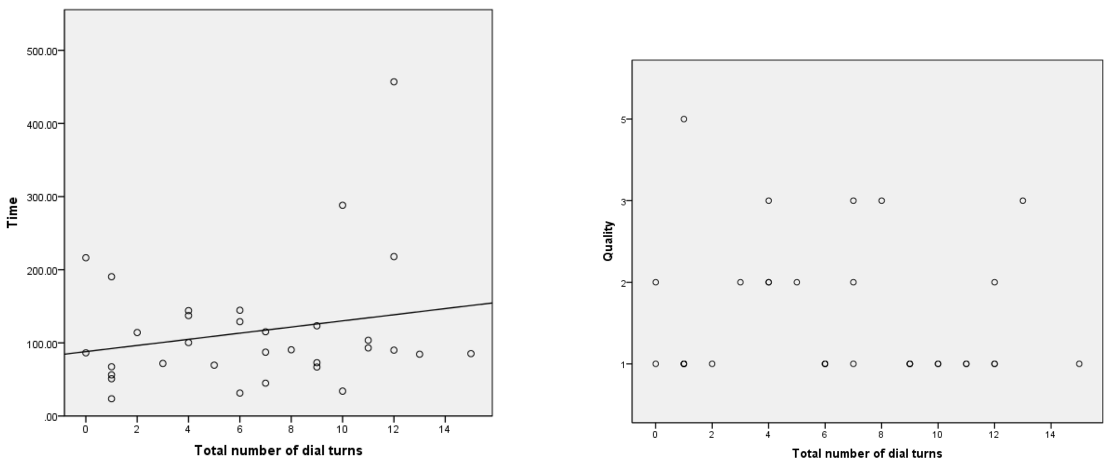


(b) Quality of Selection

Figure 38: Scatter Plots of Number of Alternates Viewed versus Performance.

used to determine the impact of familiarity of the emergency and the quality of the ranking system on the number of dial turns. In familiar scenarios, the pilots turned the dial an average of nearly two more times than they did in unfamiliar scenarios ($\mu_{fam} = 7.4$, $\mu_{unfam} = 5.5$). However, this difference was reported as non-significant, $F(1,7) = 1.865$, $p = .214$. The ranking system had very little impact on the number of times the dial was turned, $F(1,7) = .277$, $p = .615$.

Despite the lack of significant effects of the independent variables, the correlation between the number of dial turns and the pilots' performance was investigated. The Spearman's rank correlation coefficient test was used to determine this correlation. There was not a significant correlation between the number of times the dial was turned and either measure of performance. (For the correlation between number of turns and time $r_s = .148$, $p = .418$, for the correlation between number of turns and quality of selection $r_s = -.124$, $p = .497$.) Though these correlations were not significant, the direction of both correlations is as expected, more turns of the dial corresponds to higher time, but also a better decision. These correlations are seen in Figure 39.



(a) Time to Complete (b) Quality of Selection

Figure 39: Scatter Plots of Number of Times the Dial was Turned versus Performance.

5.4 Assessment Measures

Upon the conclusion of all tests runs, the pilots were asked whether they felt that the filter dial was a useful addition to the APA system. As shown in Figure 40, half of the participants responded that the addition of the dial made the system "Much better," while the other half responded that the dial made the system "Better." This feedback may be reflective of the fact that the dial may be used or ignored in any given situation. A general sentiment was expressed that a tool which can be used or ignored is always a useful addition.

The Modified Cooper-Harper for Displays was used to assess the both variations of the aid. Figure 41 shows the results both variations of the aid. Table 5.4 provides the display characteristics associated with each level of the rating scale. Every participant rated the version with the dial the same or higher than the version without the dial. One notable case was the participant who assigned both versions of the aid a score of 10, which is described as "Display is missing critical information, operator is unable to locate essential information..." This subject commented that "runway length is of critical importance and is too hard to find in the pages." All other participants rated

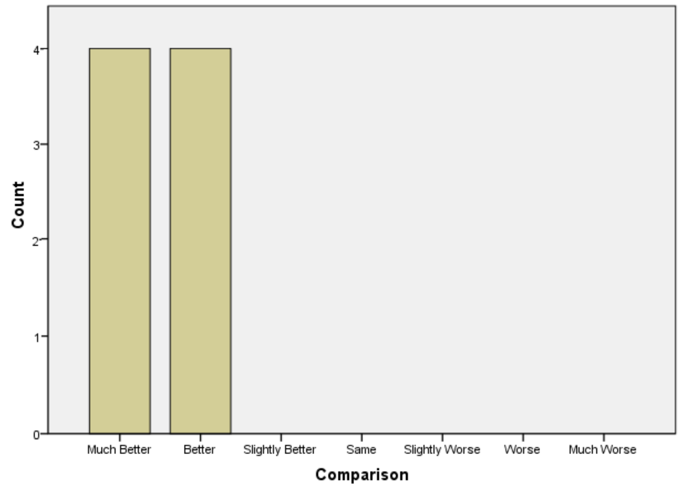


Figure 40: Comparison of APA Variations.

the variation with the dial as either “Excellent & Highly Desirable” or “Good with Negligible Deficiencies.” Only half of the participants rated the variation without the dial in either of these categories.

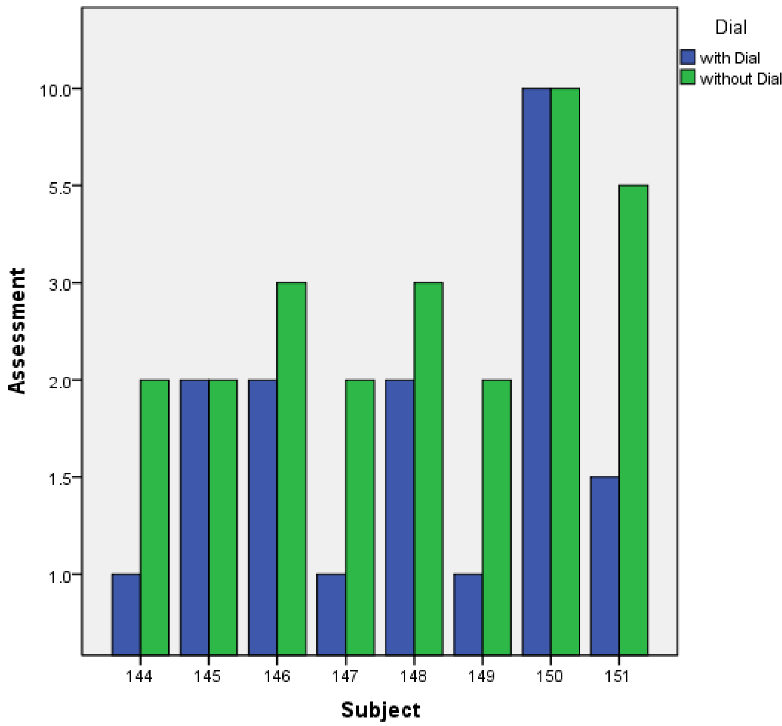


Figure 41: Modified Cooper-Harper Ratings.

Table 9: Modified Cooper-Harper Rating Scale.

1	Excellent & Highly Desirable
2	Good with Negligible Deficiencies
3	Minor but Tolerable Deficiencies
4	Moderately Objectionable Deficiencies
5	Very Objectionable Deficiencies
6,7,8	Deficiencies Require Improvement: Major Deficiencies
9,10	Mandatory Redesign: Major Deficiencies

CHAPTER VI

DISCUSSION

6.1 Applicability of Results

This experiment sought to measure pilots' performance handling emergency situations. The simulation environment was intended to provide the look and feel of an actual aircraft, but did not necessarily provide the participants with a realistic experience. All of the pilots took the experiment seriously, however, the experimental conditions could not replicate all of the experiences of an actual inflight emergency. The pilots were offered additional compensation, but this does not replicate the stress and pressures of a real emergency. Additionally, the pilots were given two training runs, which allowed them to familiarize themselves with the features available. An aid which was implemented onboard would require that pilots be trained to proficiency in order to ensure that the aid is used effectively. Conversely, emergency situations arise very infrequently, so the aid should not rely on pilots recall as they will only rarely use the aid. This importance of an intuitive interface is also mandated by the nature of the environment, which may put the pilot in an opportunistic control mode. Finally, the results presented differences which were not statistically significant, given the population size. With these effects in mind, these results should not be viewed as demonstrating differences between factors, but only an indication of possible differences.

6.2 Overall Performance

The two primary measures of performance were the time required to select an alternate landing site and the quality of the landing site which was selected. It was expected that pilots would have to make a tradeoff between these two factors. That is, to make a better decision, some pilots may more thoroughly consider their options, resulting in a long time to complete the task. On the other hand, some pilots may choose to act quickly, without investigating all options, resulting in a lower quality of decision. The dial was expected to reduce this negative correlation by reducing the number of options pilots considered. However, the results showed that the correlation between these two measures was, in fact, positive. This suggests that runs in which the pilot took longer to complete the task actually resulted in poorer landing site quality. This result may indicate that the trade-off between time and quality may not be as significant as expected. Though contrary to expectation, this may be related to the design of each scenario. One of the most important characteristics of a scenario may be the number of sites which are similar to the highest ranked site. This similarity may have made some scenarios more difficult than others, requiring more time to consider the options and less differentiation between the highest ranked sites. Another study noted that the benefit of automation decreased as time pressure was relaxed, suggesting that automation may be less beneficial in lower time pressure scenarios [19]. While that study differed in that it provided explicit time limits, a potential reduction in the benefit of automation may also contribute to this positive correlation.

6.3 Impact of Dial

The addition of the filter dial to the APA system was expected to result in improved performance, in terms of both time to complete the task and quality of landing site selected. It was thought that when pilots were given the ability to easily eliminate landing sites which were inappropriate, the process of making a selection among the reduced number of options could be completed more quickly. It was also thought that the ability to filter out options would also result in improved decisions. Though not statistically significant, the rankings of the sites selected when using the filter were somewhat better than when not using the dial.

Conversely, the time make a selection was not improved with the addition of the dial. Pilots actually took longer to select a landing site when using the APA variation with the dial. While this result is not statistically significant, it may suggest that the impact of the dial did not simplify the task as much as expected. The addition of the dial did not result in significantly fewer number of alternates for which the pilot viewed more information. There is a strong correlation between the number of alternates viewed and the time to make a selection. This may indicate that the time was not reduced by the dial because the number of alternates viewed was not reduced by the dial as expected.

6.4 Impact of other factors

A difference in performance was expected for familiar and unfamiliar scenarios. The results supported this hypothesis and showed that pilots made their selections more quickly for familiar scenarios. This supports the result found in the survey, which showed that pilots were more likely to take immediate action in a emergency with which pilots are familiar.

The effect of the addition of the dial was impacted by the familiarity of the scenario. The time to complete the task was increased by the addition of the dial in the case of familiar emergencies, while the time was decreased by the dial in the case of unfamiliar emergencies. This may suggest that the dial proved more useful for filtering out inappropriate options in unfamiliar emergencies. However, it is unclear whether this result applies generally or was impacted by the specific scenario used.

The scenarios used in the experiment were designed to be of similar difficulty. This objective was met to a large degree with the exception of the aileron control failure scenario. In this scenario, there was not one alternative which was clearly better than others. Rather, there were three options which were similarly ranked. Each of these sites had a factor that made it less than desirable such as a shorter than required runway or crosswinds. Also, given this type of failure, it is more important for the pilot to make a good decision than to act quickly. This may have lead to much internal deliberation on the part of the pilots for this scenario.

The quality of the ranking system was also expected to result in a difference in performance. However, the pilots performed very similarly in this measure regardless of whether the ranking system was optimal or non-optimal. This result may suggest that pilots were generally not misled by a poor ranking system; rather, they were able to determine the best landing sites regardless of their rank. This suggests that pilots were generally not susceptible to the over-reliance phenomenon in terms of the ranking system. Conversely, this result may be taken to imply that pilots had low confidence in the ranking system. That is, pilots were not inclined to trust that a highly ranked landing site was superior to a lower ranked site and felt that they must investigate lower ranked options. This suggests that under-reliance may be more of an issue. However, the number of landing sites for which pilots viewed more information was roughly equivalent in cases using each ranking system, which may imply that pilots investigated all options which may have been appropriate regardless of their ranking. This issue of under-reliance may also be addressed through training and a better understanding of the factors that determine the ranking of landing sites.

6.5 Processes Used

Each participant used his own method to make the APA most useful to him. However, there were some comments that shed light on how the pilots go about this task. As discussed previously, the optimality of the ranking system did not have a significant impact on the quality of the pilots' landing site selection. One point of interest is how pilots begin to narrow down the list of possible options. When the dial was available, the first step may be to adjust the filter to an appropriate settings. Interestingly, the first site for which pilots chose to view more information was the site ranked highest (by the ranking system used in the run) in almost 60% of runs.

One pilot commented that the method employed was to begin each run by assessing the alternates which were available in the scenario. This process began by looking at the airline preferred airports and the overall scores for each landing site. When an emergency occurred, the pilot was able to investigate a top choice for the emergency based on his awareness of the options available. This selection of top choice was sometimes different than the site which was ranked highest by the system provided. Another pilot commented that the method used was to determine the most important criteria for the emergency and looked up this piece of information for each of the landing sites until an acceptable option was found. This mode of operation was not well supported by the current design and may have led to much longer times to complete the task. Each of these pilots commented that they keep a separate mental ranking that may differ from the system's. This may provide one explanation for the lack of difference in landing site quality using different ranking systems.

6.6 Pilots' Comments

6.6.1 Features

After completing all test runs, the participants were asked a few questions about the features that are useful, those that are not useful, and features the APA system lacks. Most participants responded that the ability to filter out unnecessary information was one of the most useful features. Also, the ability to view more detailed information about each landing site was very useful. Pilots emphasized that a key attribute of these features was the speed with which information could be processed. The dial facilitates quickly reducing the number of options to be considered, while the MORE INFO page provides a "snapshot" of information about a particular landing site.

A number of pilots commented that the "ALL FIELDS" dial setting was not useful. Pilots said that this setting would not be useful because in any such dire emergency would not, they would not resort to the use of an APA. Contrarily, one participant commented that all information is useful and whether it is used or not should be left to the pilot in each scenario.

Pilots also provided information about features that they would like to add to the APA system. The addition of a particular filter setting was suggested by multiple participants. This setting would show all landing sites with runways of sufficient length for the aircraft. The ability to obtain critical information quickly was emphasized by pilots who suggested that more information be encoded into the graphical display. Items such as runway length and orientation, and IFR/VFR conditions were specifically recommended. Finally, pilots commented that an APA-type aid should be linked with airline dispatchers. This ability could be used in a number of ways, such as live updates of airline preferred and weighting of criteria.

6.6.2 Uses of APA

Pilots were asked to describe the types of scenarios in which an APA would be most useful and those in which it would be least useful. Most pilots expressed that the aid would be most useful in situations where there was high workload and high temporal pressure. These situations are characterized by the need to make a decision quickly. The ability to quickly access large amounts of pertinent information makes the aid a large improvement over current options. Pilots also commented that an APA would be useful in a less intense emergency, in which the aircraft is unable to reach its destination, but can proceed somewhat normally. The aid can be used more deliberately to assess all options and determine the most suitable landing site.

The situations in which the aid would not be useful were varied. One participant responded that the aid would not be useful in dire emergencies such as a fire because, "the only piece of information necessary is the nearest runway, all other data is irrelevant." However, another pilot did not identify a situation in which the aid would be least useful, but commented that, "information is always useful in formulating a plan, the more info, the better."

6.7 Updating Criteria Weights

The pilots felt that the ability to adjust the weights manually would not be useful in many circumstances, but should be included, as it may be nice to have when the emergency is less time critical. This is similar to other features of the APA which pilots felt that they may not use, but which may be included "just in case." Many participants commented that they mentally adjusted weights anyway, but would likely not take the time to alter the weights manually.

A few participants also recommended alternative methods for updating the criteria weights. A common suggestion was to allow the airline to update weights via datalink. This solution would allow the pilots to focus on the task at hand, and use the most appropriate weighting possible. Another method which was proposed was to have the pilot simply categorize each criterion as either important or not important, from this the computer would adjust the rankings accordingly. While this method loses much information, it would be quick and simple for pilots to use.

The final alternative method to update criteria weights was the use of macro-type settings. This solution would allow to pilot to select one of a number of pre-defined weightings. This may be especially useful in a familiar emergency such as an engine failure, where the emergency is somewhat predictable. However, in the event of a truly unforeseen emergency, the pilot may be left with no pre-defined weighting on which to rely. Each of these proposed methods as well as the example method may be made available in conjunction with one another, allowing the most appropriate method to be used in a given scenario.

CHAPTER VII

CONCLUSION

This thesis has sought to evaluate the efficacy of an automated path planning aid intended to help pilots plan a safe trajectory to land in the event of an in-flight emergency. This work began by soliciting input from pilots about the design of such an aid. Based on this information a prototype was designed and implemented in a cockpit simulation. This simulator was used to test the aid and gather results and further feedback from pilots.

7.1 Design Summary

The APA which was developed was based on design considerations gathered from survey responses and an analysis of the task. First, the aid which was developed had to be compatible with existing cockpit designs and should only deviate from current systems when necessary. For this reason, the aid was developed using existing Control Display Unit pages and current Navigation Display layout. Next, the aid had to be easy to use, without requiring unnecessary time and effort on the part of the pilot. This was accomplished in part by working with current systems with which pilots are already comfortable. Also, the filter dial was added to allow the pilot to quickly focus only on alternates which were appropriate for a given emergency. The final design goal was to make critical information easily accessible in a consolidated manner. The MORE INFO pages were implemented to allow pilots to find all critical data about a landing site in one location.

7.2 Experiment Summary

The primary performance metrics, time and quality, showed a positive correlation, implying that runs in which the pilot took longer to complete the task resulted in lower quality of landing site selected. This result is contrary to expectation and may indicate that the trade-off between time and quality may not be as significant as expected. The design of specific scenarios used may have influenced this result. Also, training may reduce the apparent increase in time through more efficient use of the dial.

Comparisons between the two variations of the APA showed that the addition of the dial resulted in a slight increase in quality of landing site selected, but also resulted in longer times to select a site. The dial also did not significantly reduce the number of alternates viewed, which was strongly correlated with the time metric. This may indicate that the dial did not simplify the task as much as anticipated. However, in the case of unfamiliar emergencies, the dial reduced both time to select a landing site as well as the number of solutions viewed. Every participant scored the variation with the filter dial more highly than the variation without the dial, indicating that they preferred to have the dial, despite the lack of improvement in performance. Again, after pilots have been trained to use the system more effectively, this result could change.

The aid was tested using both familiar and unfamiliar emergencies in order to understand if the APA was more useful in one type of scenario than another. Both survey and simulator results indicated that pilots are likely to act more quickly in a familiar emergency. Results also showed that the aid did not cause pilots to over rely on the ranking system. The pilots selected landing sites of similar quality regardless of how highly it ranked. This suggests that under-reliance may be an issue. However, this too may be overcome through training and a better understanding of how the ranking system works.

Pilots found the filter dial and the consolidation of information about landing sites to be very useful features. The ability to quickly and easily access critical information is one of the most important characteristic of an emergency planning aid. This design facilitated the pilots' methods of assessing each landing site throughout a flight, before an emergency has occurred. The ranking system (though not always optimal) gave the pilots aggregate scores for each site and provided a more meaningful starting place when investing the available options.

7.3 Design Recommendations

The APA that was tested was developed with a number of design considerations. The resulting aid was not, nor was it intended to be, a finished product; rather, it was a step toward a viable aid. Based on the results of this study, the aforementioned design considerations may be extended for future development. First, the aid should help pilots understand the situation and the options that are available. The requires consolidating critical information and making it quickly and easily accessible. This information must also be presented in a clearly understandable way. The MORE INFO page used in this study attempted to provide this consolidation of information. The information which is most critical may depend on situation. This may particularly include information which is abnormal, such as high volume of traffic around an airport, or runway closures. Additionally, text output may not provide the most usable means of providing this information, better use of graphical displays could allow more information to be processed more quickly by the pilot.

Ideally, the aid will also increase the amount of subjectively available time that the pilot experiences in a given situation. The results of this study suggest that reducing the amount of information that a pilot must digest may be one way of achieving this goal. The aid attempted to meet this goal by introducing the filter dial, however other methods of eliminating extraneous information may be useful as well. The ability of the pilot to provide input about which criteria are most important may be another way of filtering data. However, this must be done in a simple and intuitive manner.

In order to support pilots in a more opportunistic control mode, the aid tested attempted to merge data about each option into a single number. This reduced the data processing load by simplifying the data. Different methods of providing a snapshot of options should also be considered. Regardless of the method of consolidating information, pilots must have a better understanding of the meaning behind these numbers in order for this to be most useful. Allowing the pilot to adjust these scores by manipulating the weighting system would provide the pilot with a better understanding of the meaning, but can be a cumbersome task, only suitable in a strategic control mode.

The aid must also be flexible and support intuitive, yet effective use in a variety of emergency scenarios. This means that the aid should be useful in a scenario where every second counts, as well as a scenario with less time pressure. Similarly, the aid must support a range of control modes effectively. The aid should provide a pilot in a very opportunistic mode with only the information that is needed, which should be presented in an obvious fashion. The aid should also provide more detail, options and feedback for a pilot who has more subjectively available time. This aid should also be flexible enough to support different modes of operation, such as searching for a runway which meets a certain criterion.

Additionally, futures versions of such an aid should include ground personnel. This is a challenging task for a few reasons. First, the interaction must be made available in such a manner that the pilot may choose to interact or not as the situation allows. The interaction should also be conducted such that the pilot remains in control, and is not interrupted, but rather is able to incorporate ground resources into his or her own method.

7.4 Recommendations for Future Work

To develop a more effective aid improvements should be made to the APA design. First, a number of assumptions were made which must be addressed. Second, a few items outside the scope of this research must also be addressed. Third, responses from experiment participants have provided a number of recommendations for extending the usefulness of the aid.

In order for the APA to develop routes after failures have occurred it was assumed that the system was aware of the aircraft's post-failure dynamics and control capabilities. This may require further development and integration of sensors and aircraft monitoring systems. The post-failure performance may not even be fully deterministic, but may rely partly on a stochastic assessment of possible future states.

In order to select potential landing sites, the system requires information about all landing sites around the aircraft's path. This includes all types of airfields as well as a database of terrain information. This terrain data, as well as current weather information, will be necessary in order to construct safe landing trajectories. Finally, further research in the development of such trajectories is required in order to develop safe, time-optimal 3-dimensional paths, especially in the presence of altered dynamics and potential hazardous weather and terrain.

The current design simply calculated a time-optimal path which was presented to the pilot. A future design would likely need to take into account the severity of the emergency. In a less severe emergency, the pilot may need to take into account air traffic routes en route to the landing site. However, in more severe emergency, pilots may not only disregard traffic patterns, but also may be will to fly more aggressively and fly through some severe weather. More research is needed to better understand the best way to incorporate the pilot in the plan development portion of the task.

The comments gathered from experiment participants provided useful recommendation for additional features that may be incorporated into future versions of an

APA. First, the encoding of information into the symbology of the navigation display would further facilitate the quick and easy access to information. Specific recommendation include representation of the runway length available at each landing site and the current weather conditions (e.g. IFR or VFR).

The APA tested included four positions on the filter dial; Airline Preferred, Controlled airports, all airports, and all fields. These settings may be refined to more accurately reflect the way that pilots categorize the landing sites. Pilot comments suggest that the airline preferred setting was very useful, while the all fields setting may be removed. One setting that was missing was one which shows all landing sites whose runway is sufficiently long. Further investigation may identify other modifications to the settings which are most useful for pilots.

The current design facilitates a certain mode of operation in which the pilot is given a list of alternates. This list is sorted on some aggregate score and the pilot may choose to receive further information about any airport. However, in an alternative mode of operation, the pilot may select one or more important criteria and choose to view the data for these criteria about each airport. Under the current design, the pilot would have to sequentially choose each airport to find the single piece of information about each. Another design of the aid may better facilitate this mode of operation.

Finally, a significant direction for further development is the incorporation of personnel on the ground. This may include interactions with air traffic controllers, airline dispatchers and others. This interaction may be as simple as providing the same information to all personnel to provide a common frame of reference. It may also include allowing airline dispatchers to indicate airline preferences in a manner that is incorporated with the interface and is not a distraction from other tasks.

APPENDIX A

SAGAT QUESTIONS

A.1 SA Level 1 Questions

1. Which of the landing site has the shortest ETA?
2. How far is the nearest landing site (of any kind)?
3. How long is the nearest runway?
4. Which landing site is closest to your planned route?
5. How many landing sites are available in the APA?
6. How many weather systems are present within 160 nmi?
7. When were you expected to arrive at the original destination?
8. How far is the original destination?
9. Which of the landing sites is farthest from a weather system?
10. What is the aircraft's altitude at the end of the run?
11. How many runways of 6000+ feet were available in this scenario?
12. Which of the landing sites has the best medical services available?

A.2 SA Level 2 Questions

1. How far is the nearest airport with a runway 6000ft or longer?
2. How far is the landing site which the APA has ranked highest?

3. Which is the closest controlled airport?
4. Which is the closest airline preferred airport?
5. How many controlled airports are available in the APA?
6. What is the rank of the highest ranked airport with a runway 6000+ feet long?
7. What is the rank of the highest ranked airline preferred airport?
8. What is the rank of the highest ranked airport with medical services available that are good or excellent?
9. What is the rank of the highest ranked airport with airline maintenance available that are good or excellent?
10. What is the rank of the landing site which is closest?

A.3 SA Level 3 Questions

1. If weather were not a factor, which landing site would be ranked highest?
2. If, after turning towards the highest ranked landing site, it became unavailable, which landing site would be ranked highest?
3. Which site(s) would be eliminated if the dial were set to ALL CTRL?
4. If the flight continued its current path for 15 minutes, which landing site would be closest?
5. Which site(s) would be eliminated if the dial were set to AIRLINE PREFERRED?

APPENDIX B

BRIEFING

Thank you for your participation in the testing of the Automated Path Planning Aid. The aid is designed to provide assistance planning a diversion in the event of an on-board emergency. By testing the aid in the simulator we aim to understand what characteristics of such an aid will be most useful. This will help in further development of cockpit resources.

First you will be familiarized with the aid and other resources that will be used in the evaluations. Then you will have the opportunity to try a couple of practice runs. Once you are comfortable with the simulator, the testing will begin. You will be given nine test runs. In each run you will be the pilot not flying. Should an emergency occur, (an emergency will not occur in all runs) you will be responsible for selecting the most appropriate place to land.

After each run there will be a few questions to answer. Please do not worry about answering these questions. Some questions will seem unimportant or even impossible, but please don't even think about the questions during the simulator runs. After all nine runs have been completed you will be asked a few more questions about your experience as well some other possible aids.

In the evaluation runs, there will be two different variations of the aid that will be used. The only difference is the addition of a dial which allows you to filter out unwanted candidate routes. This will be explained in more detail shortly.

You have been told that you will have the opportunity to receive additional compensation based on your performance in the simulator. Although the exact formulation of the reward cannot be revealed, it is based on factors that would lead to a safe landing in an emergency situation, such as landing as soon as possible, without unnecessarily endangering the flight. So please treat this as you would an actual onboard emergency.

The following is an introduction to our simulator and the Automated Path Planning Aid.

Welcome to the cockpit of the next generation transport aircraft. You will find new versions of some of the controls and displays with which you are already familiar. These include the Primary Flight Display (PFD), the Navigation Display (ND), ND control panel, and the Control Display Unit (CDU).

Today, you will be the first officer of our aircraft. Your captain (not present; use your imagination) will be flying the aircraft by hand when necessary, so you will not need any other controls. Let's go through each part of the system piece by piece.

The PFD will be on the left of the screen in front of you. This will provide information about your current flight attitude such as altitude, speed, roll and pitch. As you will not be flying, this is only for reference.

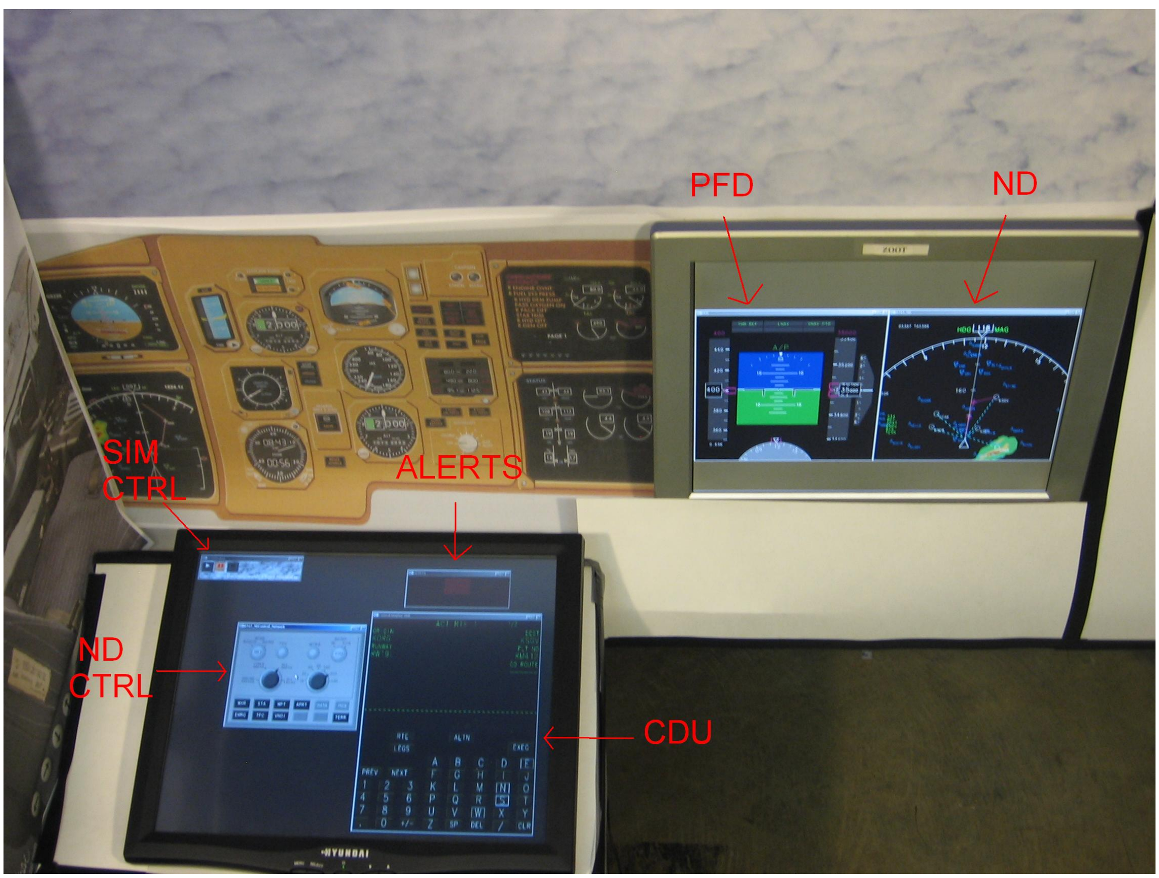


Figure 42: Simulator Displays

The ND is on the right of the screen in front of you. This provides a number of pieces of information. The display is track up, with the current plan shown in magenta. The ND control panel, on the left hand side of the center console screen, allows you to optionally show other information. There are seven options that may be turned on and off using the labeled push buttons at the bottom of the control panel.

- WX: Shows weather systems in the area. This information is updated via data link.
- STA: Shows navigation stations
- WPT: Shows all waypoints in the area
- ARPT: Shows airports, these are only airports used for large transport aircraft
- EMRG: Shows the candidate routes to alternate destinations, the routes are filtered by the APA dial (see APA dial). Routes shown as dotted gray lines.

The ND control panel also has a couple of dials. The dial on the right controls the range (in nautical miles) of the ND. In some runs the dial on the left will be the APA filter dial, in others this function will be disabled and replaced with a dummy dial. The APA filter dial (when present) allows you to filter possible divert destinations quickly. At each level, moving from right to left, more landing sites will be shown, but none will be removed. The four levels are:

- AIRLINE PRFERD: These are airports where the airline prefers that you divert. The airline will be able to perform maintenance, and route passengers on other flights.
- CTRLD ARPTS: These are airports that have control towers. There are no guarantees about the runway length, instrumentation, or emergency services available.
- ALL ARPTS: This filter shows all airports, regardless of runway length, control tower or any other factors.
- ALL FIELDS: This filter shows all potential landing sites, including any flat open space where an aircraft may potentially ditch.

Figure 43: ND Control Panel with Filter Dial.

Figure 44: ND Control Panel without Filter Dial.

The CDU provides a limited interface to some FMS functions. The RTE and LEGS pages allow you to explore the planned route. The ALTN page lists the potential divert destinations. The options shown are the same as those displayed on the ND. The PREV and NEXT buttons on the CDU may be used to change pages of the list. This list is also filtered according to the dial setting on the ND control panel. The list is order by a ranking system. This ranking system is the aids best attempt to apply weights to the criteria it is aware of. These include:

- Time to reach site
- Distance to site
- Fuel remaining upon arrival
- Length of runway
- Weather at site
- Medical services available
- Maintenance available

Please be aware that this is an automated system and the criteria weighting may not be optimal for all scenarios, but should provide a good starting point.

After selecting one of the destinations on the list, you may view more information about it by pressing MORE INFO (Line 6L). This will provide information about the airport and the values that are used by the ranking system for each of the criteria. There are two pages of information, the PREV and NEXT buttons are used to switch between them. After selecting a destination, you may press DIVERT NOW (Line 6R) to choose this option as the divert plan. (You are not required to view more information about an option in order to select it as the divert.) This will bring up the divert page, which provides a summary of the selected plan. The EXEC button at the upper right of the keypad, will also be illuminated. Pressing either the EXEC key or EXECUTE (Line 6R) on the divert page will set the selected plan as the active plan in the FMS and provide autopilot and/or flight director commands for the execution of this plan. Because you will not be executing the plan, the run will end when a plan has been executed.

During each simulator run, there is the possibility that an emergency will occur. In the event of an emergency, you will hear the captain describe the emergency. You can assume that checklists have been executed and that the captain is holding the aircraft in steady (though not necessarily straight and level) flight. You will be asked to use the tools described to select a course of action. This may be a diversion or a continuation to the original destination. Again, the run is complete when you have made this selection, you will not be doing any flying today.

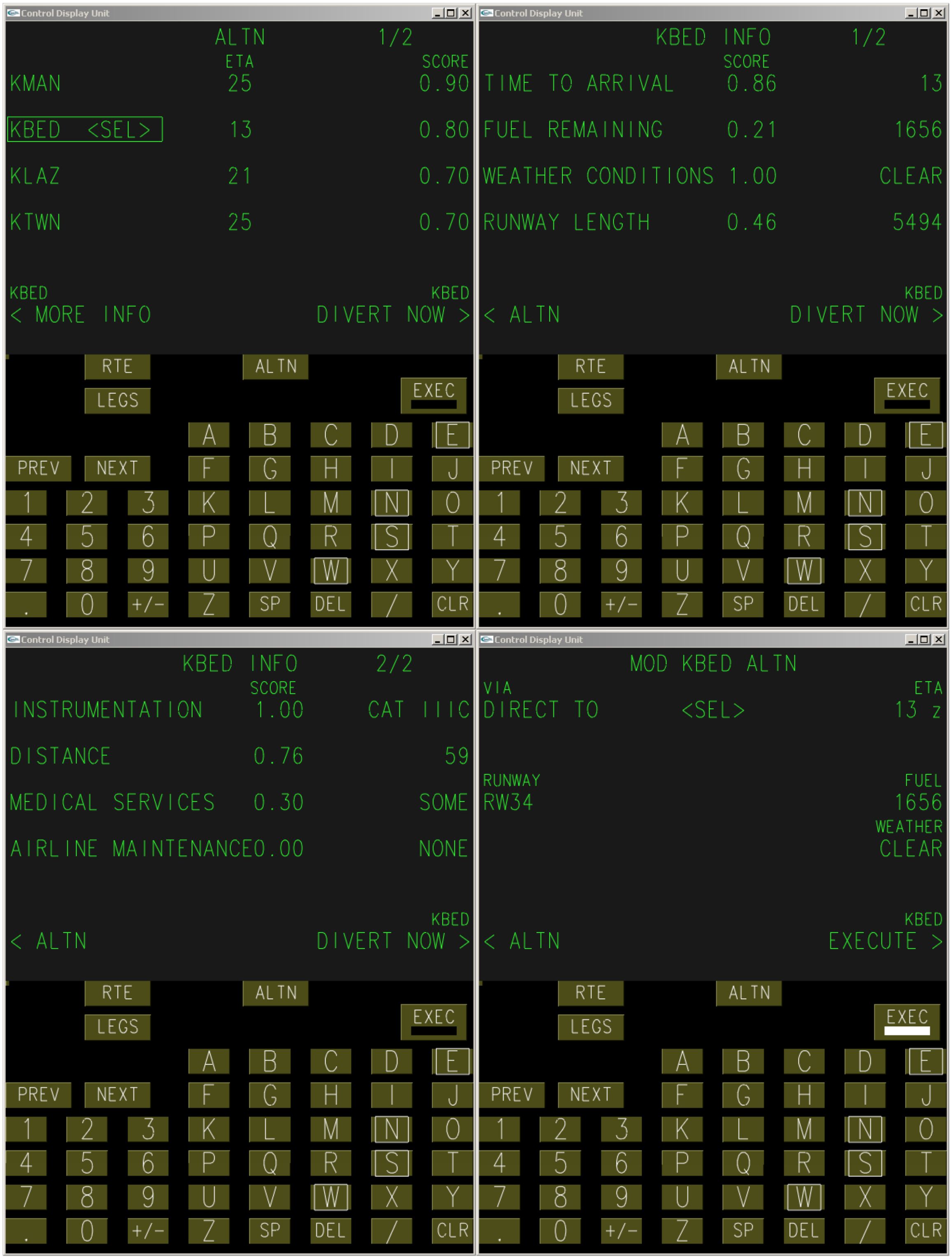


Figure 45: Upper Left: Alternates Page, Upper Right: APT INFO Page 1, Lower Left: APT INFO Page 2, Lower Right: Execute Modified Route Page.

After each run you will first be asked some questions. You do not need to be worried about correctly answering these questions, they do not impact you performance evaluation. The first few questions are about the workload you encountered during the run. Don't think too hard about these questions, just follow you first reaction. The terms used in the workload questionnaire are defined on the following page. While the terms may seem obvious, they may be used in a slightly different way in this case.

Next you will see a series of questions about the scenario you just saw. Some of the answers may seem obvious, while others may seem impossible. Please answer them as best you can and do not be concerned with trying to plan for the questions during the runs.

After completing all nine runs, you will be asked a final set of questions about your thoughts on the aid.

Are you ready for a couple of practice runs?

MENTAL DEMAND	How much mental and perceptual activity was required (e.g. thinking, deciding, calculating, remembering, searching, etc.)? Was the task easy or demanding, simple or complex, exacting or forgiving?
PHYSICAL DEMAND	How much physical activity was required (e.g. pushing, pulling, turning, controlling, activating, etc.)? Was the task easy or demanding, slow or brisk, slack or strenuous, restful or laborious?
TEMPORAL DEMAND	How much time pressure did you feel due to the rate or pace at which the task or task elements occurred? Was the pace slow and leisurely or rapid and frantic?
PERFORMANCE	How successful do you think you were in accomplishing the goals of the task set by the experimenter (or yourself)? How satisfied were you with your performance in accomplishing these goals?
EFFORT	How hard did you have to work (mentally and physically) to accomplish your level of performance?
FRUSTRATION LEVEL	How insecure, discouraged, irritated, stressed and annoyed versus secure, gratified, content, relaxed and complacent did you feel during the task?

APPENDIX C

QUESTIONNAIRE

Subject Number:

Flight experience, in hours:

What certifications do you currently hold?

What types of aircraft do you have the most experience flying?

Did you find the dial to be a useful addition?

a) Much better

b) Better

c) Slightly better

d) Same

e) Slightly worse

f) Worse

g) Much worse

What features of the APA were most useful?

What features of the APA were not useful?

What features would be useful that the APA did not have?

In what situations would the use of the APA be most and least helpful?

Are there any conditions in which you would change the route to a destination? (That is, you have chosen a destination, but would alter the automation-suggested route to it.)

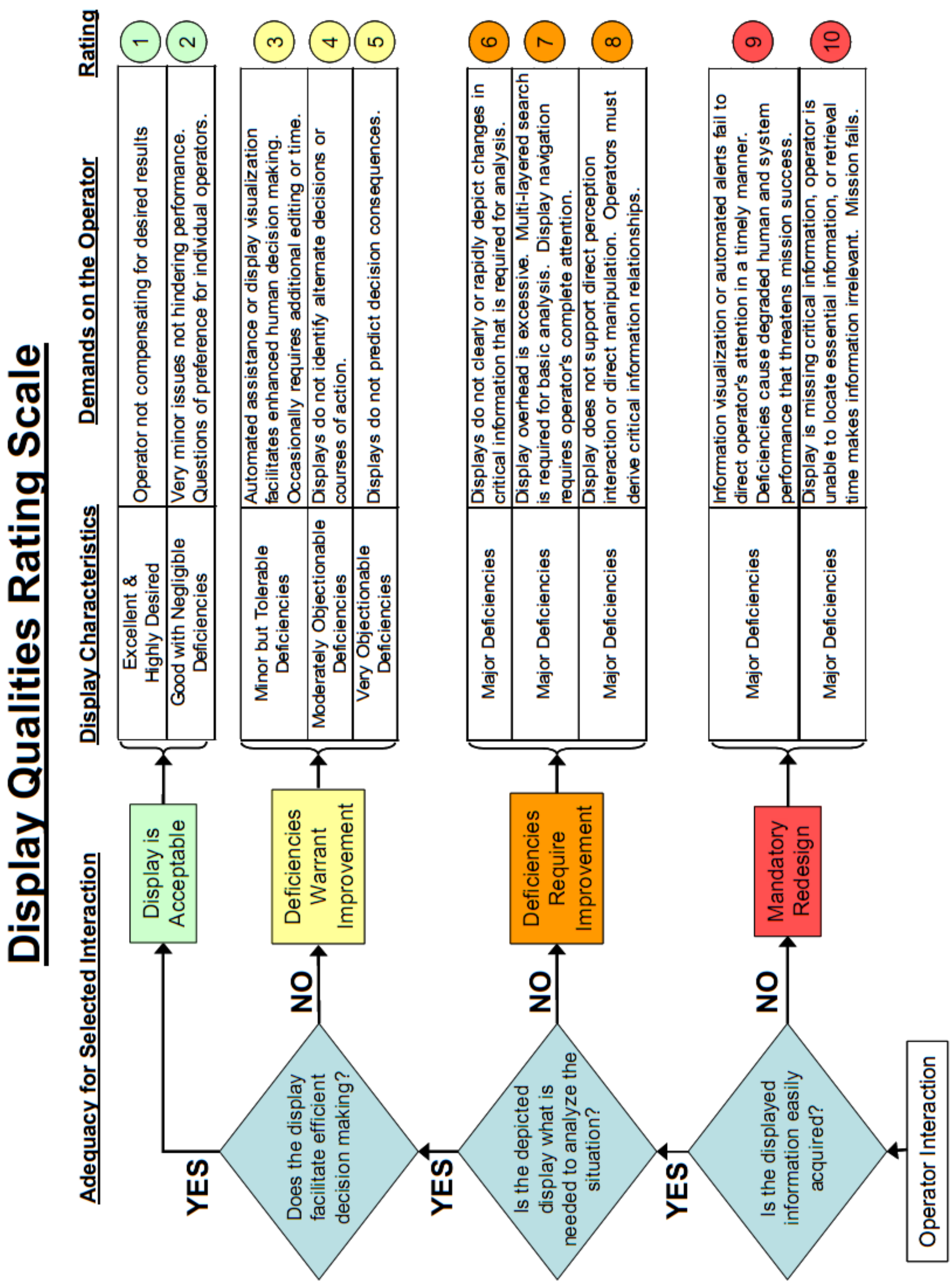


Figure 46: Modified Cooper-Harper for Displays.

APPENDIX D

DISCLOSURE OF DECEPTION

You were told prior to your participation that you would be compensated \$50 as a token of our appreciation for your support of this work. Additionally, you could receive up to \$50 based on your performance in the experiment. However, all participants will receive the full \$50 for performance. Because this study sought to investigate performance in the event of an emergency, this deception was intended to encourage participants to behave in a more motivated manner, slightly more similar to an emergency.

You have the opportunity to have your data excluded from this study if you choose. Regardless of your decision, you will be compensated a total of \$100.

Thank you for your participation and understanding the motivation for the use of deception.

REFERENCES

[1] "In-flight fire leading to collision with water. swissair transport limited mcdonnell douglas md-11 hb-iwf peggys cove, nova scotia 5 nm sw," Tech. Rep. TSB A98H0003, Transportation Safety Board of Canada, Quebec, Canada, 2 September 1998.

[2] "Statistical summary of commercial jet airplane accidents worldwide operations 1959-2008," Statistical Summary, Boeing Company, Seattle, WA, July 2009.

[3] Atkins, E. M., Portillo, I. A., and Strube, M. J., "Emergency flight planning applied to total loss of thrust," *Journal of Aircraft*, vol. 43, pp. 1205–1216, July-August 2006.

[4] Bearman, C., "Naturalistic decision making." Paper presented at the Centre for Decision Research Seminars (G. Villejoubert, Chair), Leeds University Business School, Leeds, UK., January 2004.

[5] Beyer, H. and Holtzblatt, K., *Contextual Design : A Customer-Centered Approach to Systems Designs*. Morgan Kaufmann, 1st ed., 1997.

[6] Chen, T. L. and Pritchett, A. R., "Development and evaluation of a cockpit decision-aid for emergency trajectory generation," *Journal of Aircraft*, vol. 38, pp. 935–943, September-October 2001.

[7] Cummings, M., "Automation bias in intelligent time critical decision support systems," in *AIAA 1st Intelligent Systems Technical Conference*, September 2004.

[8] Cummings, M., Meyers, K., and Scott, S. D., "Modified cooper harper evaluation tool for unmanned vehicle displays," 2006. presented at the UVS Canada Conference.

[9] Dubins, L. E., "On curves of minimal length with a constraint on average curvature, and with prescribed initial and final positions and tangents," *American Journal of Mathematics*, vol. 79, pp. 497–516, July 1957.

[10] Endsley, M. R., *Direct Measurement of Situation Awareness: Validity and Use of SAGAT*, ch. 7, pp. 147–173. Lawrence Erlbaum Associates, 2000.

[11] Endsley, M. R., "Theoretical underpinnings of situation awareness: A critical review," in *Situation Awareness Analysis and Measurement* (Endsley, M. R. and Garland, D. J., eds.), pp. 3–32, Lawrence Erlbaum Associates, 2000.

[12] Hollnagel, E., "Contextual control models of cognition," in *Human Reliability Analysis: Context and Control*, pp. 159–201, Academic Press, 1993.

[13] Hollnagel, E., "Context, cognition, and control," in *Co-operative Process Management* (Wrn, Y., ed.), pp. 27–52, Taylor & Francis, 1998.

[14] Honeywell, *Boeing 777 Flight Management System Pilot's Guide*, 2001.

[15] Human Performance Research Group, NASA Ames Research Center, *NASA TASK LOAD INDEX (TLX)*.

[16] Ippolito, C. A. and Pritchett, A. R., "Software architecture for a reconfigurable flight simulator," in *AIAA Modeling and Simulation Technologies Conference*, 2000.

[17] Jain, S. and Tsiotras, P., "Trajectory optimization using multiresolution techniques," *Journal of Guidance, Control, and Dynamics*, vol. 31, pp. 1424–1436, September-October 2008.

[18] Johannsen, G. and Rouse, W. B., "Mathematical concepts for modeling human behavior in complex man-machine systems," *Human Factors*, vol. 21, no. 6, pp. 733–747, 1979.

[19] Johnson, K. E., Kuchar, J. K., and Oman, C. M., "Experimental study of automation to support time-critical replanning decisions," in *Proceedings of the Human Factors and Ergonomics Society 46th Annual Meeting, HFES*, vol. 5, pp. 337–341, 2002.

[20] Karelahti, J., Virtanen, K., and Öström, J., "Automated generation of realistic near-optimal aircraft trajectories," *Journal of Guidance, Control, and Dynamics*, vol. 31, pp. 674–688, May-June 2008.

[21] Keinan, G., "Decision making under stress: Scanning of alternatives under controllable and uncontrollable threats," *Journal of Personality and Social Psychology*, vol. 52, no. 3, pp. 639–644, 1987.

[22] Layton, C., Smith, P. J., and McCoy, C. E., "Design of a Cooperative Problem-Solving System for En-Route Flight Planning: An Empirical Evaluation," *Human Factors*, vol. 36, no. 1, pp. 94–119, 1994.

[23] Maule, A. J. and Edland, A. C., "The effects of time pressure on human judgement and decision making," in *Decision Making: Cognitive Models and Explanation* (Ranyard, R., Crozier, W. R., and Svenson, O., eds.), Frontiers of cognitive science, pp. 189–204, Routledge, 1997.

[24] Maule, A. J., Hockey, G. R. J., and Bdzola, L., "Effects of time-pressure on decision-making under uncertainty: Changes in affective state and information processing strategy," *Acta Psychologica*, vol. 104, pp. 283–301, June 2000.

[25] Mosier, K. L. and Skitka, L. J., "Human decision makers and automated decision aids: Made for each other?," in *Automation and Human Performance: Theory and Applications* (Mouloua, M. and Parasuraman, R., eds.), pp. 201–220, Lawrence Erlbaum Associates, 1996.

[26] National Transportation Safety Board, "Accidents and Accident Rates by NTSB Classification, 1988 - 2007, 14 CFR 121." http://www.ntsb.gov/aviation/Table2.htm, 2008.

[27] Parasuraman, R. and Riley, V., "Humans and automation: Use, misuse, disuse, abuse," *Human Factors*, vol. 39, no. 2, pp. 230–253, 1997.

[28] Simpson, P. A., "Naturalistic decision making in aviation environments," Tech. Rep. DSTO-GD-0279, Air Operations Division, Aeronautical and Maritime Research Laboratory, January 2001.

[29] Watts, R., Tsiotras, P., and Johnson, E., "Pilot feedback for an automated planning aid system in the cockpit," in *IEEE/AIAA 28th Digital Avionics Systems Conference*, pp. 5.B.2–1 –5.B.2–13, October 2009.

[30] Zsambok, C. E., "Naturalistic decision making: Where are we now?," in *Naturalistic Decision Making* (Zsambok, C. E. and Klein, G., eds.), pp. 3–16, Danbury, 1997.